AFCAPS-FR-2010-0020

Identifying Leader Talent: Alternative Predictors for U. S. Air Force Junior Officer Selection and Assessment

November 2006

Michael Ingerick

Human Resources Research Organization
66 Canal Center Plaza, Suite 400
Alexandria, VA. 22314

Kenneth L. Schwartz
Johnny J. Weissmuller

AFPC/Strategic Research and Assessment Branch (SRAB)

Air Force Personnel Center
Strategic Research and Assessment
HQ AFPC/DSYX
550 C Street West, Ste 45
Randolph AFB TX 78150-4747

Approved for Public Release. Distribution Unlimited

UNCLASSIFIED

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report was cleared for release by HQ AFPC/DSYX Strategic Research and Assessment Branch and is releasable to the Defense Technical Information Center.

This report is published as received with minor grammatical corrections. The views expressed are those of the authors and not necessarily those of the United States Government, the United States Department of Defense, or the United States Air Force. In the interest of expediting publication of impartial statistical analysis of Air Force tests SRAB does not edit nor revise Contractor assessments appropriate to the private sector which do not apply within military context.

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct request for copies of this report to:

Defense Technical Information Center - http://www.dtic.mil/

Available for public release. Distribution Unlimited. Please contact AFPC/DSYX Strategic Research and Assessment with any questions or concerns with the report.

This paper has been reviewed by the Air Force Center for Applied Personnel Studies (AFCAPS) and is approved for publication. AFCAPS members include: Senior editor Dr. Thomas Carretta AFMC 711 HPW/RHCI, Dr. Lisa Mills AF/A1, Dr. Paul Ditullio AF/A1, Kenneth Schwartz HQ AFPC/DSYX, Johnny Weissmuller HQ AFPC/DSYX, and Brian Chasse HQ AFPC/DSYX.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)* 01-11-2006	**2. REPORT TYPE** Final	**3. DATES COVERED *(From - To)*** September 2004 – September 2005

4. TITLE AND SUBTITLE Identifying Leader Talent: Alternate Predictors for U.S Air Force Junior Officer Selection and Assessment	**5a. CONTRACT NUMBER** FA3089-04-F-0637 **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Michael Ingerick, Kenneth L. Schwartz, & Johnny J. Weissmuller	**5d. PROJECT NUMBER** GSF-1oF-0087J **5e. TASK NUMBER** S99-13 ALT PREDICTOR **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Human Resources Research Organization 66 Canal Center Plaza, Suite 400 Alexandria, VA. 22314	**8. PERFORMING ORGANIZATION REPORT NUMBER** GSF-1oF-0087J
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Personnel Center Strategic Research and Assessment Branch Randolph AFB TX 78150	**10. SPONSOR/MONITOR'S ACRONYM(S)** HQ AFPC/DSYX **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** AFCAPS-FR-2010-0020

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release, Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report summarizes an effort to identify the attributes needed by Air Force officers to meet the leadership requirements of the 21st century and establish the best and most efficient way to assess those attributes. Specifically, the report examined the effectiveness of current officer selection systems and practices used by the Air Force Reserve Officer Training Corps (AFROTC) and Officer Training School (OTS) programs. The report also examined alternatives to the attributes and measures currently used to seek ways to enhance assessment and selection of officer talent. (Continued on Back.)

15. SUBJECT TERMS
AFOQT, ROTC, OTS, Selection, Selection strategies

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OFPAGES**	**19a. NAME OF RESPONSIBLE PERSON** Kenneth L. Schwartz
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	U	113	**19b. TELEPHONE NUMBER** *(include area code)* 210-565-3139

Standard Form 298 (Rev. 8-98)

14. Abstract (Continued)

The findings of the report are mixed. The current selection system has areas of weakness that could be improved. However, it also appears that in some instances viable alternative procedures may not be available.. Overall the report recommends adjusting the various selection systems to one overall formal and programmatic system that can be sustained and linked with future needs and deliver a return on investment that will benefit the overall Air Force.

Executive Summary

Focus & Goals

For the Air Force and the other services, officer selection has and continues to carry significant strategic importance. Because the Air Force's junior officers represent its future, their selection constitutes one of the most critical and substantial investments the Air Force makes. Consequently, the Air Force understandably wants to know whether its investments in junior officer selection have been and will continue to be worth it. In response to this imperative, this report summarizes a systematic effort to identify (a) the attributes needed by Air Force officers to meet the leadership requirements of a 21st Century expeditionary force, and (b) how best and most efficiently to assess these attributes. Specifically, this effort aimed to answer the following questions:

- How effectively are the selection systems (and practices) currently in use by the Air Force Reserve Officer Training Corps (AFROTC) and Officer Training School (OTS), in particular the Air Force Officer Qualifying Test (AFOQT), meeting the Air Force's needs and performance imperatives?
- What alternatives to the attributes (e.g., general mental ability) and measures (e.g., AFOQT, GPA) currently employed by these systems carry the greatest potential to enhance the assessment and selection of officer talent?
- Taking into account these findings, what specifically would an Air Force officer selection and assessment –toolkit" look like?

Method

To answer these questions, we comprehensively reviewed (a) Air Force strategic plans, initiatives, and leadership doctrine (e.g., *Personnel Strategic Plan, FY 2004-2009*; *Transformation Flight Plan 2004*), and (b) relevant theory and empirical research (e.g., meta-analyses, individual research studies, technical reports) on military and organizational leadership, supplemented by expert judgments (as needed). In addition, we interviewed and/or consulted Air Force representatives involved in officer selection from AFROTC, Air Force Recruiting Service (AFRS), the Air Force Personnel Center (AFPC), and the Air Force Directorate of Personnel (AF/DP).

Findings

Key findings, as they relate to the first two questions (above), include:

- Evidence is limited that the current selection systems, and their components, are meaningfully contributing to the Air Force's ability to meet current and future competency (performance) requirements. For example, there presently is no demonstrable, systematic process for linking officer selection to competency (performance) requirements, as spelled out in Air Force doctrine.

- Differences in how the different commissioning sources define officer ―quality" necessitate the Air Force to make tradeoffs that could affect its return on investment (ROI) in officer selection. These differences are reflected in the different measures AFROTC and OTS use and the relative weight assigned to each when making selection decisions.
- In some instances, it is not always clear exactly what attributes the current selection systems are measuring, or more importantly the degree to which they are measurably linked to the Air Force's competency (performance) requirements (e.g., the Relative Standing Score [RSS] employed by AFROTC). In addition, AFROTC and OTS employ several unstructured assessments (e.g., letters of recommendation) that could be improved to increase their ROI and to facilitate standardization across the different commissioning sources.
- Existing research conducted by the Air Force indicates that replacing the subtests comprising the AFOQT Academic Aptitude (AFOQT-AA) composite, which measures general mental ability (GMA), with a comparable standardized test, specifically the SAT Reasoning Test,[1] would not significantly affect the prediction of officer success. Further, recent revisions to the SAT (e.g., inclusion of subtest assessing written communication skills) could serve to enhance this prediction.
- A number of additional attributes not directly assessed by the current selection systems, such as higher-level cognitive and cross-functional skills (e.g., social skills), could significantly add to the prediction of officer success (commissioning and post-commissioning).

Recommendations

From these findings, several actionable recommendations were generated. The following recommendations, organized by their focus, emerged as most critical:

Selection System Strategy

- Develop and implement a competency (performance) model and align officer selection systems accordingly, so that performance requirements and imperatives – to include officer retention – measurably drive selection. More long-term, ensure there is a formal, programmatic system in place for sustaining and maintaining linkages between officer selection systems (and their components) and future changes to competency (performance) requirements and other critical imperatives.
- Implement a sufficiently comprehensive criterion assessment system that enables the measurement of desired results (e.g., officer performance and retention), so that the success of officer selection systems and their components can be effectively assessed.
- Investigate and take advantage of ways in which current (or proposed) selection tools could be integrated with recruitment and training to compliment and increase the Air Force's collective ROI.

---

[1] The new SAT Reasoning Test, formerly the Scholastic Achievement Test and Scholastic Aptitude Test (SAT), consists of three sections: (a) Critical Reading, (b) Mathematics, and (c) Writing. The test went operational in 2005.

Selection System Content

- Partner with AFRS and AFROTC (and USAFA) to standardize the specification and conceptualization of predictor attributes in use for officer selection.
- Maximize coverage of predictor attributes not currently assessed, such as cross-functional skills (e.g., social skills), whose addition could significantly increment the prediction of officer success and promote Force diversity.

Selection System Method, Design, and Implementation

- Meaningfully and measurably separate assessment of officership from the assessment of technical expertise (or potential).
- Partner with AFRS and AFROTC to structure measures used in officer selection to promote and facilitate standardization across commissioning sources.
- Partner with AFRS to systematically review and evaluate specific measures used in OTS officer selections with limited ROI (e.g., letters of recommendation).
- Investigate implementation issues and related concerns (e.g., faking and coaching) that could preclude the use of the Self-Description Inventory (SDI+) as an officer selection and classification tool.
- Systematically study and investigate the possibility of using the new SAT Reasoning Test in place of the GMA-specific subtests in current AFOQT.
- Investigate possibilities for a joint-service junior officer selection battery.

To support the implementation of these recommendations, several research areas were identified.

THIS PAGE WAS INTENTIONALLY LEFT BLANK.

CONTENTS

Introduction....................1
Focus and Goals....................2
Overview....................2

How the Air Force Selects Junior Officers: An Overview....................3
Air Force Reserve Officer Training Corps (AFROTC)....................3
Mission and Program Overview....................3
Current Selection System....................5
Officer Training School (OTS)....................10
Mission and Program Overview....................10
Current Selection System....................11

Making the Grade: Evaluating Current Officer Selection Systems....................13
Step 1: Defining Selection System Success—Formulation of Effectiveness Criteria....................13
How to Define Officer Selection System Success? A Conceptual Framework....................14
What Are the Forces and Trends Driving 21st Century Air Force Leadership and Selection Requirements?—The Causes....................16
What Do These Forces and Imperatives Mean for Air Force Officer Selection?—The Effects....................16
What Makes an Air Force Officer Selection System Successful? A Review and Integration....................24
Step 2: Identifying Critical Gaps in Current Officer Selection Systems....................24
Selection System Strategy....................24
Content....................27
Methods, Design, and Implementation....................32
Air Force Officer Qualifying Test (AFOQT)....................33
Step 3: Recommendations for Closing Critical Gaps....................39
Short-Term....................39
Selection System Strategy....................39
Selection System Content....................40
Selection System Method, Design, and Implementation....................41
Long-Term....................43
Selection System Strategy....................43
Selection System Content....................43
Selection System Method, Design, and Implementation....................43

Towards an Officer Assessment “Toolkit” for the 21st Century: Identification and Evaluation of Alternative Predictors and Assessment Methods..45
Step 1: Identifying Predictor Attributes to Target for Selection.. 45
What We Did 45
Findings..46
Step 2: Identifying and Evaluating Assessment Methods.. 55
What We Did 55
Findings..56
Step 3: Recommendations for a 21st Century Officer Assessment ―Toolkit”......................... 64
Officer Recruitment and Commissioning (Training)...64
Officer Selection 66
Implications of Implementing the Proposed ―Toolkit”..68

Moving Ahead: A Summary of What is Needed and Next Steps 71
What is Needed... 71
Next Steps 71
An Implementation Road Map..77
Recommended Research to Support Road Map 77

References..81

Appendix A: Demographic Composition of Air Force (Active Component) Officer Accessions by Source of Commissioning for FYs 2001-2004 .... A-1

Appendix B: AFOQT and SAT-ACT Equivalency Chart B-1

IDENTIFYING LEADER TALENT: ALTERNATIVE PREDICTORS FOR U.S. AIR FORCE JUNIOR OFFICER SELECTION AND ASSESSMENT

Effective leadership constitutes a vital and inimitable resource for organizations (Hogan & Kaiser, 2005; Yukl, 2005; Zaccaro & Klimoski, 2001). The decisions leaders make and how they perform can exert a substantial impact on a range of individual, team, and organizational outcomes – explaining upwards of 25% of the variance in various indices of success (cf. Barrick, Day, Lord, & Alexander, 1991; Day & Lord, 1988; Harter, Schmidt, & Hayes, 2002; Judge & Piccolo, 2004; Judge, Piccolo, & Illies, 2004). Like many organizations, the Air Force invests significant resources and time in assessing, selecting, and developing its future leaders. Unlike private and public sector organizations, however, the Air Force cannot simply “buy” leader talent as needed. For the Air Force and the other Services, commissioning represents its sole means for securing leader talent, both short-term and long-term, for the officer ranks. Therefore, the effectiveness by which the Air Force assesses and selects junior officers for commissioning carries significant strategic importance, more so than for most organizations. As a result, the Air Force understandably wants to know whether its investments in officer selection have been and will continue to be worth it.

As Figure 1 illustrates, the Air Force’s assessment and selection of junior officers takes place in context. Since the 1990s, there have been significant forces and trends challenging the Air Force that hold implications for officer quality. For example, on the “demand” side, the Air Force is undergoing major transformation efforts, as the U.S. military shifts from an industrial-age, Cold War structured force; to an information-age, Post-Cold War force. At the same time,

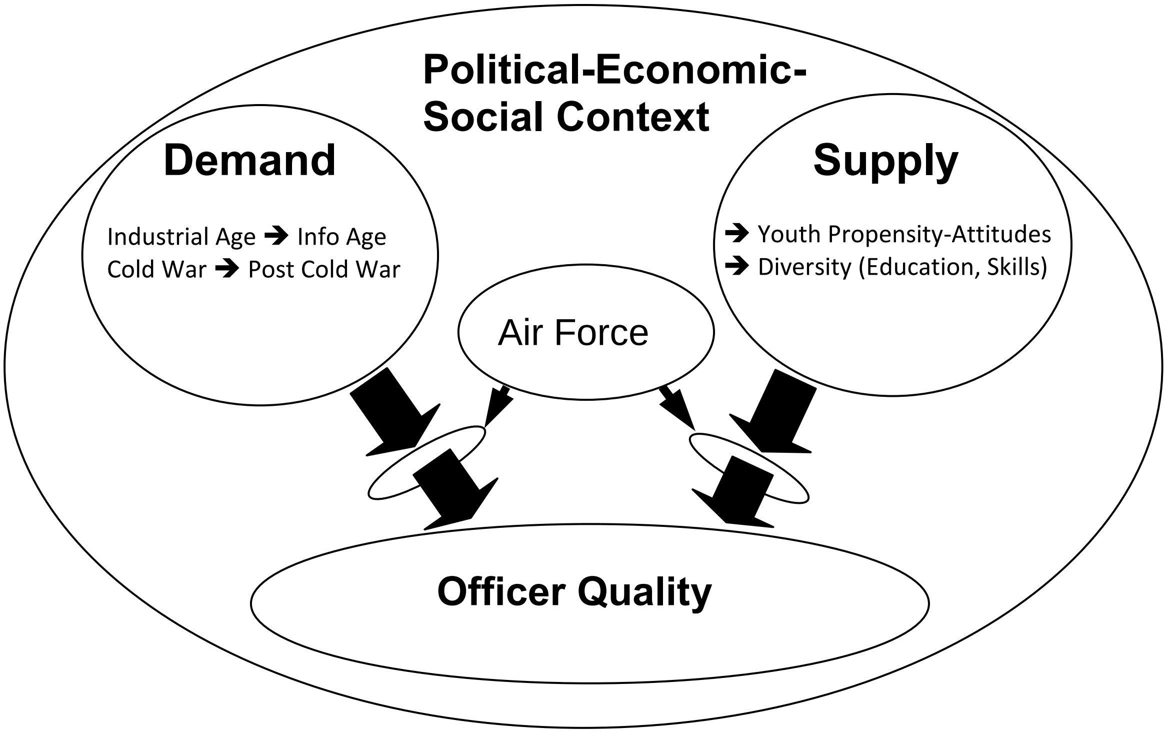


Figure 1. Forces and Trends Impacting Air Force Officer Quality

on the ―supply” side, there have been concomitant changes in the military-age youth population, such as shifts in youth’s propensity and attitudes towards military service and growing diversity in the kinds of educational experiences and attributes targeted youth bring to the Air Force. Taken together, these trends are affecting who will become officers and what they must do for the Air Force to be successful as a 21st Century global expeditionary force.

The Air Force is strongly committed, both tactically and strategically, to meeting the challenges and imperatives created by these changes not only to maintain but more importantly to enhance officer quality. As evidence of this commitment, there have been several recent, substantive changes to the Air Force Officer Qualifying Test (AFOQT), the Air Force’s flagship test for assessing junior officers for commissioning (cf. Weissmuller, Schwartz, Shore, & Gould, 2004). Nevertheless, many of the platforms and methods for assessing and selecting Air Force junior officers have generally remained unchanged for a number of years. Consequently, there is a need to comprehensively review and evaluate the Air Force’s current officer selection platforms and practices in light of these imperatives.

Focus and Goals

In response to this need, this report summarizes a systematic effort to identify those attributes Air Force officers need to meet the leadership requirements of a 21st Century expeditionary force and how best to assess them. More specifically, the goals of this research were: (a) to review the selection systems (and practices) currently in use by the Air Force Reserve Officer Training Corps (AFROTC) and Officer Training School (OTS), in particular the AFOQT; (b) to identify and compare alternatives to the predictor attributes (e.g., general mental ability [GMA]) and measures (e.g., AFOQT, GPA) currently in use that have the potential to enhance the assessment and selection of officer talent; and (c) generate actionable recommendations for an Air Force officer selection and assessment ―toolkit.”

Since all Air Force officers are leaders first and technical specialists second, this report does not address the attributes and issues specific to the selection and/or categorization of (a) rated officers (Pilots, Combat Systems Operators, and Air Battle Managers) or (b) non-line officers (judge advocates, chaplains, and medical services officers). Similarly, because AFROTC and OTS represent the largest suppliers of commissioned officers to the Air Force and are functionally separate from the U.S. Air Force Academy (USAFA), how the USAFA assesses and selects officer candidates was not included in our evaluation. Nevertheless, since the platforms and practices employed by AFROTC and OTS are comparable to those used in the Air Force Academy and the fundamental challenges and needs of the Air Force are the same across the different commissioning sources, the report’s findings and recommendations should be applicable.

Overview

This report is organized as follows. First, we summarize the current systems and practices in use at AFROTC and OTS for selecting officer candidates. Second, we conduct a strategically-focused review of these systems to identify potential gaps and to make recommendations for closing them. Third, we identify and compare alternative predictors and assessment approaches with the potential to significantly enhance current selection systems, proposing a ―toolkit" that can be employed across the recruitment-selection-commissioning of officer candidates. Finally, we conclude by identifying the most pressing priorities for implementation and research.

How the Air Force Selects Junior Officers: An Overview

Starting in February 1997, AFROTC and OTS were realigned under the newly created Air Force Officer Accession and Training Schools (AFOATS) to ensure coordinated leadership and policy direction in the recruitment, selection, and training of Air Force junior officers. As Figure 2 illustrates, about two thirds of newly commissioned officers begin their careers upon completion of one of these two programs.[2] The following sections provide an overview of AFROTC and OTS, their missions, and a summary of how each source currently selects officer candidates for training.

Air Force Reserve Officer Training Corps (AFROTC)

Mission and Program Overview

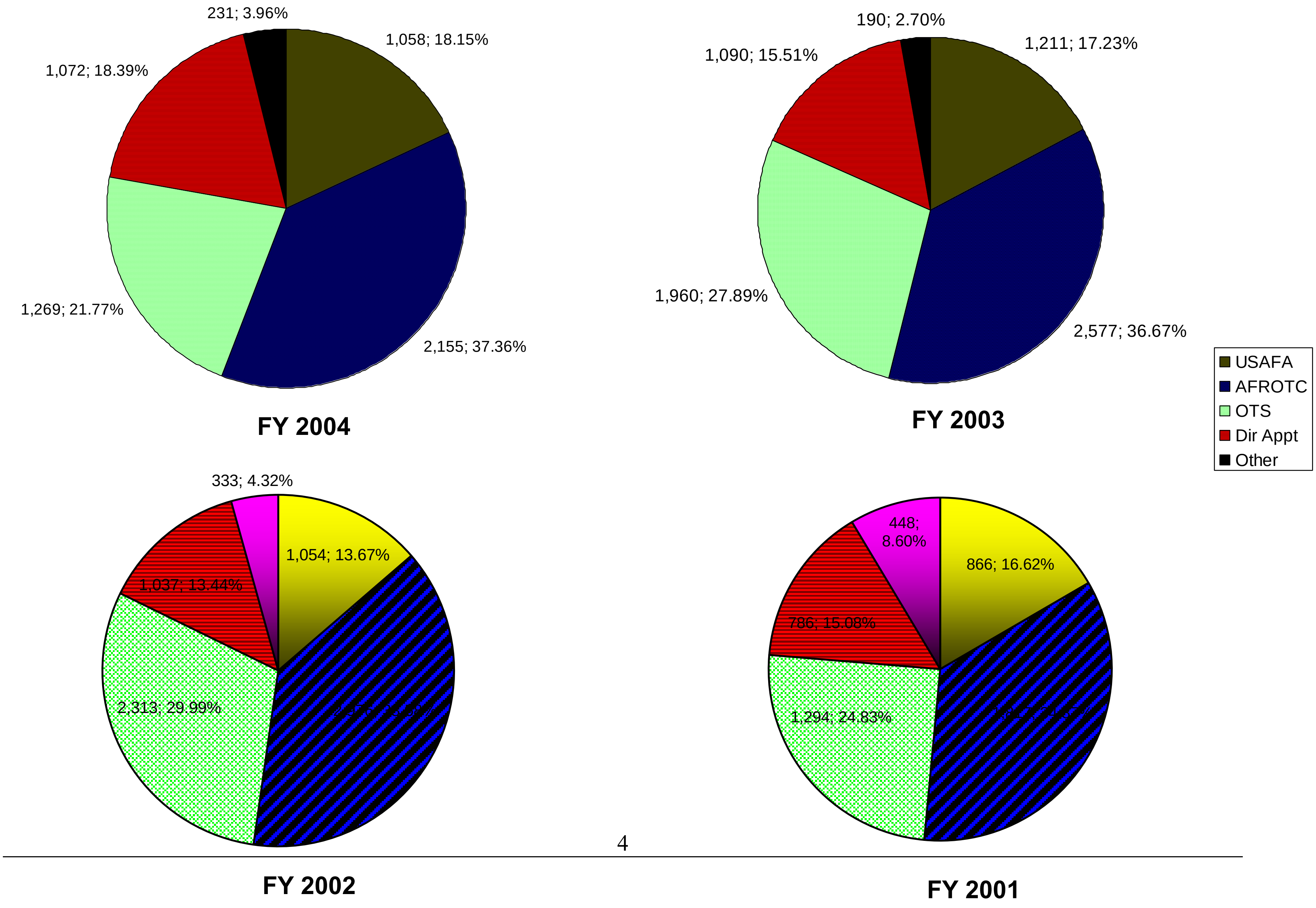
Mission. AFROTC's mission is to recruit, educate and commission future leaders for the Air Force through college campus programs based on Air Force requirements. Headquartered at Maxwell AFB (Montgomery, AL), the AFROTC program manages and administers detachments or units at 144 college campuses throughout the U.S. and Puerto Rico (as of 2003). Students from schools near AFROTC host institutions can attend classes through more than 850 separate cross-town enrollment programs or consortium agreements. Recent AFROTC enrollments, on average, have been around 17,000 to 19,000. During the past four FYs (2001-2004), the number of officers completing AFROTC has ranged from 1,817 (FY 2001) to 2,976 (FY 2002) – about a 10-20% completion rate.

Program Overview. The AFROTC training program consists of an academic phase and a field training phase. The academic phase is divided into two two-year courses: the General Military Course (GMC) and the Professional Officer Course (POC). The GMC consists of a weekly Aerospace Studies class (one hour) and a Leadership Lab (one-two hours). The GMC emphasizes leadership skills and professional military standards, customs, and procedures. In addition to formal coursework, there are optional programs where cadets can participate in specialized off-campus learning experiences (e.g., professional development and training, base visits, and so on). Cadets accepted into the POC continue to attend weekly Aerospace studies classes and Leadership Labs. POC coursework emphasizes leadership, communication skills, and military and international security studies. In addition, cadets apply what they learned in the GMC and in the field. For example, POC cadets conduct the GMC Leadership Labs and assist in

---

[2] Figures from *Population Representation in the Military Services* (OSD, 2002, 2003, 2004, 2005). Data on recent (Active Component) officer accessions by gender and race/ethnicity can be found in Appendix A.

Figure 2. Air Force (Active Component) Officer Accessions by Source of Commissioning (FYs 2001-2004)



FY 2004

FY 2003

FY 2002

FY 2001

managing their respective unit's cadet corps. Depending on when a cadet applies to AFROTC and their status at entry – as a civilian (includes members of the Air Force Reserve, Air National Guard, prior service personnel, and those from the other Services) or an active duty airman – the academic phase is either two or four years in length.[3] Cadets applying as civilians prior to starting college or with three academic years remaining in their degree program must participate in and complete both the GMC and the POC.[4] All cadets that apply as active duty airmen or as civilians with two academic years remaining in their degree program must complete the POC.

The field training phase is either four or six weeks in length. Cadets complete field training at an assigned Air Force base prior to starting the POC. Field training introduces cadets to the Air Expeditionary Force (AEF) concept and emphasizes military leadership and discipline. It consists of classroom learning, exercises and activities targeting aircraft and aircrew orientation, Air Force professional development, marksmanship training, officer training, and physical fitness training. As with the academic phase, the length of the field training depends on when a cadet enters AFROTC and his/her entry status – as a civilian or as an active duty airman. Cadets entering AFROTC as civilians who have successfully fulfilled requirements for the GMC and been accepted into the POC complete a four-week field training course. Cadets applying as active duty airmen likewise complete the four-week course. Cadets who did not participate in and complete the GMC are required to complete the six-week course.

All cadets attend AFROTC classes and activities along with other college courses. Cadets in the AFROTC program normally receive elective academic credit for fulfilling these requirements. At each host institution, AFROTC has the status of an academic department. Instructors are active duty Air Force officers – most have at least a master's degree and are usually accorded the academic rank of assistant professor. The unit commander has an academic rank of professor.

For all cadets, successful completion of a four-year accredited degree program and AFROTC leads to a commission as a second lieutenant.

Current Selection System

Eligibility. Applications to AFROTC can come from currently active duty Air Force personnel and civilians (includes members of the Air Force Reserve, Air National Guard, prior service personnel, and those from the other Services). To apply, prospective cadets must be enrolled full-time at an accredited college or university that hosts an AFROTC unit or a college (or junior college) having a cross-town agreement or consortium arrangement with an accredited institution that does. In addition, *all* prospective non-rated, line officer candidates (active duty or civilian) must meet the minimum aptitude, educational, and physical fitness requirements. Tables 1 and 2 compare eligibility requirements and officer selection processes and practices for ROTC and OTS.

---

[3] In limited cases where the Air Force has critical needs for select academic majors, there is a one-year program.

[4] Depending on the experiences they bring to AFROTC, cadets with prior service may have some or all of the GMC waived by their detachment (or unit) commander.

Table 1. Overview Comparison of AFROTC and OTS Eligibility Requirements

Details	**AFROTC-PSP**		**OTS**	
	Enlisted	Civilian	Enlisted	Civilian
Who Can Apply	Students (a) enrolled full-time at an accredited college or university that hosts an AFROTC unit or have an arrangement with an accredited institution that does, AND (b) accepted into one of the airman commissioning programs (AECP, ASCP, or SOAR), OR (c) seeking entry into the POC-ERP.	Students (a) enrolled full-time at an accredited college or university that hosts an AFROTC unit or have an arrangement with an accredited institution that does, AND (b) have satisfactorily completed the GMC, OR (c) have completed the first two years of their academic degree program.	(a) Graduates of a regionally accredited college (or university), OR (b) college seniors who will be available to depart for training within 270 days, AND (c) have one year or more of continuous active service and at least one year on station.	(a) Graduates of a regionally accredited college (or university), OR (b) college seniors who will be available to depart for training within 365 days.
Minimum Standards	Age (18-35)[1] AFOQT-Q (≥ 10) AFOQT-V (≥ 15) Pass PFT Pass Physical Pass Unit Screening/ Processing[2]	Aerospace Studies Grades (no less than C-) Age (18-35)[1] AFOQT-Q (≥ 10) AFOQT-V (≥ 15) Passed all Leadership Labs Pass PFT Pass Physical Pass Unit Screening/ Processing[2]	Academic Degree[3] Age (18-34) AFOQT-Q (≥ 10) AFOQT-V (≥ 15) Good Moral Character GPA (≥ 3.0)[4] Pass Physical	Academic Degree[3] Age (18-34) AFOQT-Q (≥ 10) AFOQT-V (≥ 15) Good Moral Character GPA (≥ 3.0)[4] Pass Physical

Notes. [1] Scholarship applicants must be less than 31 years old as of December 31 of the year they are commissioned. Non-rated, line officer candidates must be commissioned by age 30 (or waiverable up to age 35).

[2] Prior to nominating them, each unit must screen applicants to determine their motivation toward pursuing an Air Force commission. The screening includes, but is not limited to drug use and alcohol abuse, civil involvements, and academic goals.

[3] Non-Rated applicants must have earned a qualifying academic degree for the panel (Critical Technical, Technical, and TDSP) to which they are applying.

[4] For Critical-Technical and Technical applicants, there is no minimum GPA requirement. For Non-Technical applicants with a total AFOQT composite score (AA + Q + V) of 150 or greater, the GPA minimum of 3.0 may be waived.

Table 2. ***Overview Comparison of AFROTC and OTS Officer Selection Platforms and Processes***

	AFROTC-PSP		**OTS**	
Details	Enlisted	Civilian	Enlisted	Civilian
Platform & Structure	Board	Board	Board divided into sub-panels (Critical Technical; Technical Degree Sponsorship Program; Technical; and Non-Technical) and for select panels further divided by career field or academic discipline.[1]	Board divided into sub-panels (Critical Technical; Technical Degree Sponsorship Program; Technical; and Non-Technical) and for select panels further divided by career field or academic discipline.[1]
Evaluation Model	—Whole Person”	—Whole Person”	—Whole Person”	—Whole Person”
Who Evaluates	Board scores are computed objectively used pre-defined formula.	Board scores are computed objectively used pre-defined formula.	Each panel is composed of 5 colonels (or colonel selects) who independently review and score each application.	Each panel is composed of 5 colonels (or colonel selects) who independently review and score each application.
Factors Evaluated	(a) Relative Standing Score (RSS) (b) Cumulative College GPA (c) PFT Score (d) AFOQT-AA or SAT/ACT Score[2]	(a) Relative Standing Score (RSS) (b) Cumulative College GPA (c) PFT Score (d) AFOQT-AA or SAT/ACT Score[2]	(a) Education/Aptitude (b) Experience (c) Potential/Adaptability	(a) Education/Aptitude (b) Experience (c) Potential/Adaptability
Weighting	RSS (50%), GPA (20%), PFT (15%), AFOQT-AA or SAT/ACT (15%)	RSS (50%), GPA (20%), PFT (15%), AFOQT-AA or SAT/ACT (15%)	All three factors weighted equally (1/3rd).	All three factors weighted equally (1/3rd).
Selection Based On	Total board score used to rank applicants by relative order of merit. Quality cut lines established, taking into account Air Force end strength requirements. Those that exceed cut line are awarded enrollment allocation to POC.	Total board score used to rank applicants by relative order of merit. Quality cut lines established, taking into account Air Force end strength requirements. Those that exceed cut line are awarded enrollment allocation to POC.	Total board score used to rank applicants by relative order of merit. Quality cut lines established, taking into account Air Force end strength requirements. Those that exceed cut line are selected.	Total board score used to rank applicants by relative order of merit. Quality cut lines established, taking into account Air Force end strength requirements. Those that exceed cut line are selected.

Notes. [1] Within each panel, applications from enlisted airmen and civilians are reviewed and ranked separately.

[2] When computing applicants' total board score, either AFOQT-AA or total SAT/ACT scores are used, depending on which is highest.

Platform, Structure, and Processes. Of the three Air Force officer commissioning sources, AFROTC is the most challenging to describe. Generally speaking, for many cadets, AFROTC follows a multiple hurdle selection model with a probationary period. Cadets are initially selected from multiple entry points, with slightly differing criteria depending on their status (civilian or active duty Air Force) and whether they receive a scholarship (and from which scholarship program). Once selected and contracted, cadets complete the GMC and/or the first two years of their academic degree program. Upon successful fulfillment of these requirements, cadets are then eligible to become officer candidates and access to the POC, which all cadets must complete for commissioning. The POC Selection Process (PSP) governs cadets' selection to the POC.

For civilians seeking enrollment in AFROTC, there are typically two initial points of entry: (a) a scholarship program (e.g., Four-Year or Three-Year Scholarship) or (b) direct enrollment in GMC (for non-scholarship cadets).[5] Since the majority of AFROTC cadets do not start with a scholarship (about 85-90%), most enter directly by enrolling in the GMC. Contracting with GMC takes place on a ―fully qualified" basis – provided a prospective cadet is enrolled in an accredited institution and meets the minimum eligibility requirements, s/he is qualified to enroll and participate in the GMC. Depending on the program, scholarship selection similarly takes place on a fully qualified basis (e.g., Express Scholarship and Minority Institution Scholarship Programs) or by board (e.g., Four-Year or Three-Year Scholarship Programs). For those programs using a selection board, the process is comparable to that employed by the USAFA.[6]

Each AFROTC Scholarship Selection Board consists of panels of three senior Air Force officers (colonels or colonel selects). Prior to making their evaluations, all AFROTC Scholarship Selection Board members receive standardized instructions and complete some practice (or ―mock") assessments. Upon completion of these instructions and practice assessments, panel members individually rate each applicant using the ―whole person" concept. Specifically, each applicant is evaluated on the following three general factors, with measures (or indicators) of each in parentheses: (a) Academics/Aptitude (SAT/ACT scores, high school GPA/rank, high school transcripts); (b) Experience/Leadership Potential (employment history, extracurricular and community service activities, personal interview with an active duty senior Air Force officer, written responses to application questions); and (c) Commitment/Fitness for Military Service (physical fitness, personal interview with an active duty senior Air Force officer, written responses to application questions).[7] After the board evaluates all applications and any significant scoring discrepancies (i.e., across raters) are resolved, each applicant is assigned a board (or total) score. Using their board scores, applicants are ranked based on their relative order of merit. Scholarships are then offered in board rank order, based on the needs of the Air Force.

---

[5] Civilians who decide to pursue a commission through ROTC with two years remaining on their academic degree program are selected under the PSP.

[6] Except that, in contrast to the USAFA, ROTC applicants are not required to obtain a congressional nomination.

[7] Instructions for conducting ROTC Scholarship Selection interviews are contained in AFI 36-2249, *Evaluating USAF Academy (USAFA) Candidates and Air Force Reserve Officer Training Corps (AFROTC) College Scholarship Applicants*. Interviewers complete a standardized evaluation form (AF Form 4060, *USAFA Candidate Evaluation/AFROTC College Scholarship Evaluation*), rating each applicant on items covering seven dimensions (*Character/Core Values*; *Self-Confidence*; *Human Relations*; *Planning & Organizing*; *Communication Skills*; *Leadership*; *Motivation Toward Air Force*) using a five-point scale. Values associated with each point on the scale are then added across items (and dimensions) to obtain an overall recommendation (1 = *Not recommended* to 5 = *Outstanding-exceptional applicant*). Additionally, interviewers are provided space for recording specific observations on one or more of these dimensions.

For active duty Air Force personnel, there are similarly multiple points of entry into AFROTC. At present, there are four: (a) Airman Education and Commissioning Program (AECP), (b) Airman Scholarship and Commissioning Program (ASCP), (c) Scholarships for Outstanding Airmen to AFROTC (SOAR), and (d) Professional Officer Course-Early Release Program (POC-ERP).[8] Overall, the qualifications for these programs are the same. Where they differ is in the prospective candidates' current academic progress and preferred enlisted status while pursuing their commission and completing their degree (e.g., those in AECP maintain their active duty status while pursuing commissioning, whereas those in the other programs are allowed to separate from active duty until completing AFROTC). All follow a selection board process comparable to the AFROTC Scholarship Selection Boards. Consistent with AFROTC Scholarship Boards, Enlisted Boards consist of panels of three senior officers (colonels or colonel selects; at least one of whom is an officer assigned to an AFROTC unit). Prior to making evaluations, Enlisted Board members receive standardized instructions and go through some practice (or ―mock") assessments. Upon completing these instructions, members of each panel independently rate each applicant using the ―whole person" concept – rating each applicant on the following three dimensions, based on the measures (or indicators) in parentheses: (a) Academic/Aptitude (AFOQT scores, high school or college GPA, college major, degrees); (b) Military Performance/Leadership Potential (unit commander's ratings, enlisted performance reviews scope/level of responsibility, disciplinary actions); and (c) Physical Fitness (Physical Fitness Test [PFT] scores). After the board evaluates all applications and any significant scoring discrepancies are resolved, each applicant is assigned a board (or total) score. To be eligible for selection, applicants must receive a board score of 55 or greater – 100 is the maximum. When computing the board score, Military Performance/Leadership Potential receives the greatest weight (50%), followed by Academics/Aptitude (35%) and Physical Fitness (15%). Once computed, applicants are ordered by relative order of merit based on their board-score. Taking into consideration Air Force end strength requirements, a quality cut line is established, which may vary from board-to-board. Applicants exceeding the cut line are eligible to enroll in AFROTC.

To become an officer candidate and gain access to the POC, all cadets (enlisted or civilian; scholarship or non-scholarship) must be processed through the PSP. As with the scholarship and enlisted selection processes, the PSP follows a board process.[9] PSP Boards, which meet annually in the Spring, are managed by and convened at AFROTC headquarters (HQ AFROTC). Cadets that meet minimum eligibility requirements (e.g., AFOQT scores, cumulative college GPA, PFT score) and initial screening by their respective units (for drug use and alcohol abuse, civic involvements, academic goals, and so on) are nominated to the PSP Boards. While PSP Boards follow the ―whole person" concept, in contrast to other board selection processes, they do not numerically score applicants but review the ratings made by the applicants' AFROTC unit and their standardized test scores, which are then weighted to compute an overall board score from which applicants are ordered by relative merit.

---

[8] Although they are administratively assigned to an AFROTC detachment (or unit), AECP participants attend OTS upon completion of their degree to be commissioned.

[9] Except for active duty airmen participating in AECP. AECP participants are awarded an enrollment allocation to POC on a fully qualified basis.

The factors that go into computing this overall board score (or Order of Merit-OM) are: (a) Relative Standing Score (RSS), a value calculated by comparing the unit commander's ranking of applicant relative to other cadets in his/her peer group competing for selection to class size; (b) cumulative college GPA; (c) Physical Fitness Test (PFT) score; and (d) either AFOQT-AA score or total SAT/ACT score, depending on which is highest.[10] To compute the board score (or OM), each factor is weighted as follows: RSS (50%), cumulative GPA (20%), PFT (15%), and AFOQT-AA or SAT/ACT (15%). Once all scores are computed, applicants are ranked by their relative order of merit. To meet the needs of the Air Force, other factors (e.g., academic major; foreign language ability) may be considered when preparing these rankings. As with other board processes, the PSP Board establishes a quality cut line, taking into account Air Force end strength requirements. Applicants exceeding the cut line are awarded enrollment allocations in the POC.

Officer Training School (OTS)

Mission and Program Overview

Mission. Same as the AFROTC, the mission of OTS is to train and commission quality officers for the Air Force. Established in 1959 and located at Maxwell AFB, OTS serves as the ―flexible partner" of Air Force commissioning programs. As the flexible partner, OTS increases or decreases its production (as needed) to meet the Air Force's needs for officer talent. Consistent with this role, the number of officers trained and commissioned by OTS can vary year-to-year in response to fluctuations between projected Air Force end strength requirements and actual USAFA and AFROTC officer accessions. For the past four FYs (2001-2004), the number of junior officers commissioned under OTS has varied from 1,269 (FY 2004) to 2,313 (FY 2002).[11] Of these, upwards of 85% have been non-rated, line officers.

Program Overview. OTS manages and administers two officer training programs: Basic Officer Training (BOT) and Commissioned Officer Training (COT). Both BOT and COT share the same goal – to instill high standards of conduct and provide officer candidates with the essential military knowledge and skills needed for effective performance – and cover the same subject areas. Where the two programs differ is that BOT targets rated and non-rated, line officers, while COT services non-rated, non-line officers (judge advocates, chaplains, and medical services officers). Additionally, COT is a more compact program lasting four weeks, whereas BOT is twelve weeks in length.

To ensure graduates possess the knowledge and skills needed to be effective officers, the OTS curriculum covers (a) leadership (e.g., leadership styles, management principles, time management, problem-solving, goal setting, and ethics), (b) professional knowledge and military standards (e.g., dress and appearance, customs and courtesies, military law, pay and allowances, officer evaluations and career progressions, and workplace issues), (c) communication skills (e.g., oral communication, written communication, and listening skills), and (d) military and international security studies (e.g., military history and doctrine, nature and laws of armed conflict, U.S. internal and foreign policies, national security issues, and Air Force strategy). While OTS candidates receive instruction and guidance in these areas concurrently, the first half of training emphasizes teambuilding, followership, and knowledge acquisition with the second half emphasizing leadership application. To facilitate their development, OTS candidates are exposed to an integrative blend of instructional methods, including formal lectures, readings, guided discussions, classroom exercises, field leadership exercises, and after hours training activities. One of the highlights of OTS is the *Leadership Reaction Course*, a field exercise using a specialized obstacle course where small groups of candidates practice handling stress in situations that test their ability to reason quickly and lead effectively to get the group through. OTS concludes with *Vigilant Warrior*, the capstone field leadership assessment exercise, in which candidates demonstrate their ability to integrate and apply the communication and leadership skills learned throughout the course.

---

[10] AFOATSI 36-2013, *AFROTC POC, Pilot and Navigator Allocations Management* contains the conversion table used for equating AFOQT-AA and total SAT/ACT scores (p. 5) to determine which of these standardized scores is highest. A copy of this table can be found in Appendix B.

[11] For additional data on recent OTS accessions, see Appendix A.

As with AFROTC, successful completion of BOT leads to a commission as a second lieutenant. Depending on their professional credentials, COT graduates are typically awarded ranks ranging from second lieutenant to lieutenant colonel.

Current Selection System

Eligibility. As with AFROTC, both active duty Air Force personnel and civilians (includes members of the Air Force Reserve, Air National Guard, prior service personnel, and those from the other Services) are eligible to apply for commissioning through OTS. In contrast to AFROTC, however, applicants to OTS must be either (a) a graduate of a regionally accredited college (or university), or (b) a college senior who is available to depart for training within 365 days (if civilian) or 270 days (if active duty Air Force). To be eligible to apply, *all* prospective non-rated, line officer candidates (active duty or civilian) must meet the minimum aptitude, educational, and physical fitness requirements summarized in Table 1.

Platform, Structure, and Processes. Prospective OTS candidates are selected using a board process comparable to that used in officer promotions. Responsibility for managing and administering these selection boards falls under the Air Force Recruiting Service (AFRS). At present, because AFRS and OTS are functionally separate and have different missions, OTS has limited input into how AFRS structures and manages its officer recruitment and selection processes.

AFRS convenes six selection boards throughout the year, four Critical Boards (quarterly) and two Non-Critical Boards (bi-annually). The Critical boards are divided into five panels: (a) Pilot, (b) Navigator (Combat Systems Operator), (c) Air Battle Manager, (d) Critical Technical (Engineering and Meteorology), and (e) Technical Degree Sponsorship Program (TDSP). The Non-Critical boards are divided into two panels (a) Technical (Computer Science, Math, Chemistry, Physics, Architecture, and Computer or Electrical Engineering), and (b) Non-

Technical (all other non-rated classified degrees or academic disciplines). For the Rated (Pilot, Navigator (Combat Systems Operator), and Air Battle Manager), Critical Technical, and Technical panels, applications are further subdivided by career field and/or academic discipline (e.g., Technical includes the various engineering disciplines, meteorology, and others). Non-Technical applications are *not* divided by potential career field or academic discipline. Additionally, because they bring differing qualifications and experiences, applications for active duty airmen and civilians are reviewed separately to ensure each is rated fairly against their respective peers.

Each panel is composed of five colonels (or colonel-selects) who review all applications assigned to that panel. Prior to reviewing applications, a formal charge with specific instructions is provided to all panel members. In addition, members are briefed on the scoring process and the scoring split resolution process, since any difference in scores greater than 1.5 must be resolved. After receiving these instructions, panel members independently score each applicant using the ―whole-person” concept. Specifically, and as summarized in Table 2, applicants are rated on the following three factors, with measures (or indicators) of each in parentheses:

- Education/Aptitude (academic major; AFOQT scores; GPA; and college transcripts);
- Experience (letters of recommendation; employment history; military experience/performance; scope/level of leadership responsibility; awards/honors/recognition received; community service or base involvement activities; and athletics/skills/hobbies); and
- Potential/Adaptability (evaluation of interviewing officer; letters of recommendation; communication skills; and law violations).

For each factor, applicants are rated on a scale of 1 to 10 with 1/10-point increments on each factor – with 7.0 and below being *Below Average* and 9.1-10.0 being *Absolutely Superior*. As mentioned previously, should panel members differ in their ratings by more than 1.5 points for any given applicant, his/her application is returned to the panel for rescoring. After each panel member has scored that panel’s applications and resolved any splits, each applicant’s scores are summed to derive a board score (i.e., total score). To be eligible for further consideration, applicants must receive a board score of 30 or greater – 50 is the maximum. When computing the board score, each factor is weighted equally, counting one-third towards the total score. This scoring process is identical to that used by Air Force’s officer promotion boards.

Once all scoring splits have been resolved and board scores computed, qualifying applicants are ordered by relative order of merit, based on their board score – this is done for each panel and sub-panel (where applicable). Cut-lines, which determine who is ultimately selected, are then set based on officer quality considerations and Air Force end strength requirements for that specific career field (or area). The results of the board process are then forwarded to the Air Education and Training Command (AETC/CC) for final approval.

Making the Grade: Evaluating Current Officer Selection Systems

To start, we conducted a focused, strategically-oriented evaluation of the selection systems, in particular the AFOQT, in use by AFROTC and OTS. Completing this evaluation consisted of three, iterative steps (a) formulating a set of effectiveness criteria (i.e., how would the Air Force define selection system –success”), grounded in the Air Force’s needs and performance imperatives, to guide and focus the evaluation process; (b) evaluating the existing selection systems at AFROTC and OTS, in particular the AFOQT, against these criteria to identify critical gaps; and (c) generating specific, actionable recommendations for addressing these gaps. Figure 3 summarizes this process.

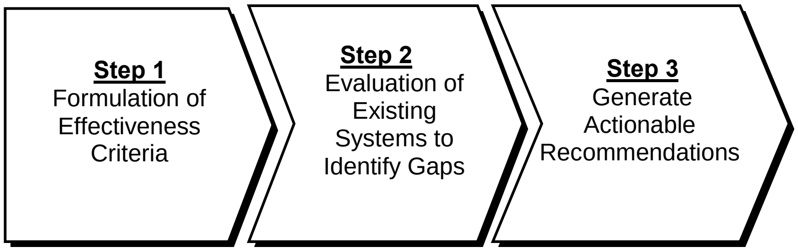

Step 1
Formulation of Effectiveness Criteria
Step 2
Evaluation of Existing Systems to Identify Gaps
Step 3
Generate Actionable Recommendations

Figure 3. Evaluation Steps

Step 1: Defining Selection System Success – Formulation of Effectiveness Criteria

Effective organizational leadership constitutes a vital and inimitable resource, impacting a wide range of individual, team, and organizational outcomes (Barrick et al., 1991; Day & Lord, 1988; Harter et al., 2002; Judge & Piccolo, 2004; Judge et al., 2004). Consequently, how an organization chooses the right people – those with the right qualifications to be leaders, at the right time and place – is critically and strategically important. Since it develops and grows its leaders internally, this is especially true for the Air Force. Unlike non-military organizations, the recruitment and selection of young men and women with the requisite attributes to become successful officers represents the Air Force’s sole means of acquiring leader talent. As a result, the Air Force understandably wants to know whether its investments in officer selection make sense given the emerging and future challenges it faces in shaping and sustaining a 21st Century force.

While we generally know a great deal about how to design and develop good selection platforms and assessments – ones that are reliable, valid, and psychometrically sound – such technical criteria, in and of themselves, do *not* ensure that a selection system will achieve its desired and expected impact (Guion, 1998; Higgs, Papper, & Carr, 2000). Ultimately, to be effective a selection system, and its individual components, must enable the Air Force to successfully meet its strategic needs and performance imperatives – a position consistent with the Air Force’s new Personnel Vision and emphasis on strategic human capital management (see U.S. Air Force, 2003, 2004b). Research consistently supports this proposition demonstrating that organizations whose HR systems and processes are strongly aligned with their needs and performance imperatives experience greater effectiveness (cf. Higgs et al., 2000; Wright & Boswell, 2002).

How to Define Officer Selection System Success? A Conceptual Framework

Figure 3 provides a conceptual framework for answering this fundamentally important question. As the Figure illustrates, leader competence and performance (i.e., what leaders do) can exert a significant influence on several individual, unit, and organizational outcomes – ranging from individual, unit, and organizational effectiveness (e.g., mission success) to outcomes associated with more intangible, value-based benefits (e.g., airmen morale and culture). For an organization to enjoy these outcomes its leaders must possess – or demonstrate the potential to develop – the attributes (i.e., aptitudes and abilities, skills, and personal qualities) needed to successfully enact the competencies and performance required to produce them. Ultimately, these performance requirements and attributes needed for success flow from imperatives (e.g., cognitive, informational, social), which are driven by forces operating within the larger political-economic-social environment (see Figure 1).

Similarly, and equally as important, each of these imperatives carries implications for how leaders are recruited, selected, assessed, and trained. Specifically, as it applies to leader selection, these imperatives carry implications for selection system (a) strategy (i.e., *why* select and for what?), (b) content (i.e., *what* attributes to assess?), and (c) methods, design, and implementation (i.e., *how* best to assess and use these assessments to make selection decisions?). Generally speaking, selection systems that are measurably aligned with these imperatives, and similarly with other HR systems, are more likely to produce the leaders an organization needs to be successful.

To understand these imperatives and their implications, we comprehensively reviewed (a) Air Force strategic plans, posture statements, initiatives, and leadership doctrine (e.g., U.S. Air Force, 1996, 2003, 2004a, 2004b, 2005), (b) strategic plans and initiatives produced by the Department of Defense (DoD) and the Joint Chiefs of Staff (e.g., U.S. Department of Defense, 2001; U.S. Joint Chiefs of Staff, 2000, 2003), and (c) research on emerging trends and future force requirements impacting the Air Force and/or the other Services (e.g., Ford, Campbell, Campbell, Knapp, & Walker, 2000; Galway, Buddin, Thirtle, Ellis, & Mele, 2005; Horey & Fallesen, 2004; Horey et al., 2004; Sager, Russell, Campbell, & Ford, 2004; Weaver, 2001; Wong, Bliese, & McGurk, 2003; Zaccaro, Klimoski, & Boyce, 1999). In addition, we consulted with Air Force decision-makers representing AFPC and AF/DP. These reviews and consultations served to answer two central questions:

- What are the forces and trends driving 21st Century Air Force leadership and officer selection requirements (―causes”)?
- What do these forces and the imperatives they engender mean for Air Force officer selection (―effects”)?

In the following sections, we summarize our findings to these questions, which formed the basis for the effectiveness criteria – how the Air Force would define selection system ―success” – subsequently used to evaluate the current selection systems (and their components).

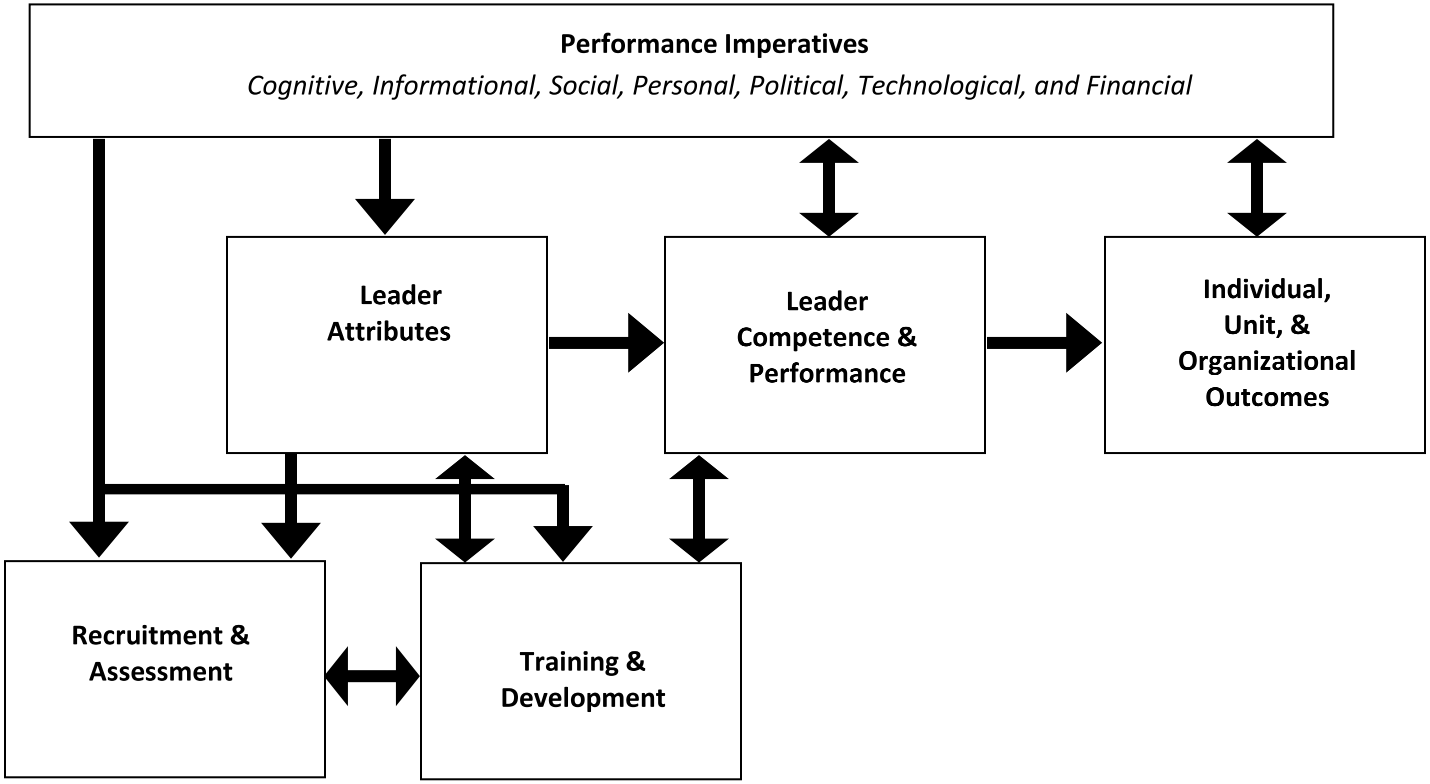

Performance Imperatives
Cognitive, Informational, Social, Personal, Political, Technological, and Financial
Leader Attributes
Leader Competence & Performance
Individual, Unit, & Organizational Outcomes
Recruitment & Assessment
Training & Development

Figure 4. Conceptual Framework for Understanding Air Force Officer Leadership and Selection Requirements (Adapted from Zaccaro, 2001)

What Are the Forces and Trends Driving 21st Century Air Force Leadership and Selection Requirements? – The Causes

Generally speaking, the causal forces impacting 21st Century Air Force leadership requirements (see Figure 1) can be organized into two overarching categories, Demand and Supply. *Demand* represents those forces influencing the Air Force's mission (i.e., what needs to get done and how), whereas *supply* represents those forces shaping (a) the kinds of individuals available to become future officers, and (b) the kinds of individuals that these officers will lead. Table 3 summarizes the most critical forces associated with each, along with illustrative trends, that have been identified by previous research and/or are reflected in existing Air Force strategic plans and initiatives (e.g., Total Force Development).

What Do These Forces and Imperatives Mean for Air Force Officer Selection? – The Effects

As with the other Services, these forces are transforming Air Force strategy, its structure and organizational culture, and what it means to be an effective junior officer. As many civilian organizations have discovered (cf. Higgs et al., 2000; Ilgen & Pulakos, 1999; Pearlman, 2003; Pearlman & Barney, 2000), these forces are changing, and in some cases, creating new performance imperatives that are influencing how the Air Force selects current and future officers. As mentioned previously, these imperatives carry specific implications for selection system (a) strategy, (b) content, and (c) methods, design, and implementation. Consistent with this, our findings are organized around these components.

Imperatives and Implications Impacting Selection System Strategy. As with other organizations, the forces and challenges the Air Force faces have created different kinds of strategic choices when conceiving selection systems and have raised the stakes surrounding conventional choices. Chief among these effects is the increasingly visible and important role selection plays in achieving organizational goals and objectives (Guion, 1998; Higgs et al., 2000; Pearlman, 2003; Pearlman & Barney, 2000). Generally speaking, the imperatives impacting selection system strategy carry implications (a) for *why* an organization selects (i.e., what is a selection system being designed to accomplish?), and (b) for *what* the organization selects on (i.e., what is the selection system being designed to predict?). Table 4 summarizes the performance imperatives affecting Air Force selection system strategy and what they mean for officer selection.

Perhaps most important for Air Force selection system strategy is what these systems should predict. Recently, the Air Force, as have the other Services, has formulated a leadership- or officer-focused competency model to answer this question.[12] Briefly, this model consists of three meta-competencies – Leading the Institution, Leading People/Teams, and Personal Leadership – within which are more specific competencies, arranged at three different levels of leadership – tactical, operational, and strategic. As with competency models in general, this model intends to (a) provide a systematic mapping of the performance requirements critical to

---

[12] Developed by the Air Force Senior Leader Management Office (AFLSMO), a full description of the model and its components can be found in AFDD 1-1, *Leadership and Force Development*.

Table 3. Overview of the Forces and Imperatives Driving 21st Century Air Force Leadership and Selection Requirements

Nature	**Forces**	**Trends and Initiatives**
Demand	Geopolitical-Military	• Enemies using new and evolving tactics (e.g., terrorism, information warfare, etc.); • High global operational and personnel tempo; • New and greater range of operations (e.g., homeland security, peacekeeping, MOOTW, etc.) requiring flexible mix of capabilities and personnel; • Increase in joint, coalition, and multinational operations; • New and flexible —communities of service” (e.g., civilian-contractor-military); • Evolving personnel and training doctrine and other Force Transformation initiatives (within and outside of the Air Force); • Need to incorporate and leverage operational lessons learned —real time” into training and other processes.
	Technological	• Transformation from an industrial age to an information age force requiring use of new equipment, platforms, and weapons capabilities; • Increased focus on technology as a force multiplier; • Increased volume, speed, and complexity of information; • Greater use of digitized and virtual, instead of face-to-face, communication and collaboration.
	Economic	• Uncertain and continued pressure on military budget and priorities (e.g., operational needs versus strategic R&D); • Continued Force Shaping and other initiatives to re-size enlisted and officer ranks; • Systematic efforts to transform the way the Air Force conducts business; • Economic pressures that challenge the Air Force’s ability to effectively promote and support Airmen quality-of-life.
	Socio-Cultural/ Socio-Political	• Increased focus and emphasis on internal and external customer service; • Evolving U.S. foreign policy, U.S. and international community perceptions of military operations; • Greater focus on public opinions and expectations by civilian decision-makers; • Instant and increased media presence during operations.
Supply	Demographic	• Decrease (from 1990s) in youth propensity and attitudes towards military service (and those of their influencers); • Changes and increased diversity in youth and Airmen characteristics (e.g., race/ethnicity, educational experiences, personal qualities, etc.); • Projected workforce shortages in select occupations (or career fields).
	Economic	• Increased competition from civilian sector in recruiting and retaining quality Airmen (officers and enlisted).

Table 4. Summary of the Performance Imperatives and Implications Impacting Selection System Strategy

Performance Imperative	**Rationale**	**Implications**	
		Why Select?	For *What*?
Fundamental need to select junior officers to effectively meet the performance (competency) requirements of a 21st Century expeditionary force.	Demand-side forces (e.g., Geo-Political, Technological, etc.), coupled with Force-wide transformation initiatives in response to these forces, are potentially re-making the performance (competency) requirements of junior officers – increasing the strategic importance of selecting the right people.	• Selection system should exert (and demonstrate) Force-wide impact, selecting the officers with the right qualifications to facilitate success and Force transformation.	• Select for attributes predictive of wide range of role- or behaviorally-based and strategically driven leadership competencies and core values. • Because of these expanded performance (competency) requirements, assess greater number and kinds of attributes.
Growing need for flexible, versatile Airmen (i.e., generalists) to meet Total Force and emerging operational capabilities.	Newer and greater range of operations, coupled with Force Shaping and other Total Force initiatives.	• Selection system (and its components) should facilitate the achievement of Total Force goals. • Officer selection should be meaningfully separate from selection (or classification) of candidates to technical specialties.	• Select for performance requirements and —core” (or broader) attributes associated with generalists (i.e., adaptability, continuous learning, etc.).
Increased pressure on recruitment and retention of junior officers.	Increasing competition for prospective candidates (or current officers), specifically for particular specialties (e.g., Critical Technical), coupled with Air Force's need to attract and retain intellectual capital.	• Selection system components (i.e., content, methods, design, etc.) should contribute to the accomplishment of recruitment and retention goals.	• Select for performance *and* retention. • Select for person-environment —fit”, specifically (a) person-role (leader), and (b) person-organization fit.
Increased potential for misalignment (or gaps) between prospective candidates' attributes/qualifications and performance requirements – for officer and/or selected technical specialty.	Workplace forecasts and trends (e.g., increased diversity in youth population characteristics), coupled with increasing and growing complexity in Air Force performance requirements.	• Selection system components (i.e., content, methods, design, etc.) should support and facilitate training and development (both for those candidates selected and those not selected).	• Select for —core” (or broad) attributes strongly associated with learning (i.e., ability and motivation to learn, adaptability, core self-evaluation, etc.) or development potential.

future Air Force leaders, and (b) systematically instruct the Air Force's Force Development efforts, ensuring the coordinated design and implementation of the full spectrum of HR systems (e.g., selection, training, promotion). As Table 5 shows, the competencies comprising the Air Force's model have been identified by one or more of the other Services, even though they may differ in their level of specificity, terminology, and how each defines a ―competency." Similarly, many of these competencies, or similar ones, can be found in existing leader and supervisory performance models (e.g., Avolio & Bass, 2002; Borman & Brush, 1993; Conway, 2000; Yukl, 1990), as well as more general models of performance (e.g., Campbell, McCloy, Oppler, & Sager, 1993; Hedge & Teachout, 1986; Motowidlo, 2003).

To ensure that the Air Force's model provides sufficient coverage of those competencies critical to 21st Century military leadership, we compared the model to those constructed from comparable efforts to specify Future Force officer competency (performance) requirements (e.g., Ford et al., 2000; Fallesen & Reichard, 2005; Garstka, 2003; Hedge, Borman, Bruskiewicz, & Bourne, 2004; Horey & Fallesen, 2004; Horey et al., 2004; Pulakos, Arad, Donovan, & Plamondon, 2000; Zaccaro, 1999; Zaccaro et al., 1999). Table 6 provides a representative sampling of competencies identified by these efforts. Differences in terminology notwithstanding, and consistent with the previous comparison, we found substantial conceptual overlap in the competencies identified by the Air Force and those identified by other researchers. Where there were differences, they tended to be in the inclusion and/or more explicit coverage of competencies related to *Cultural Adaptability-Competence* and *Information Management*, both of which are expected to be relevant to current and future Air Force officers. In particular, Cultural Adaptability-Competence is emerging as extremely important. There are several reasons for this, including (a) the changing and increased diversity of Air Force personnel, (b) the greater number of joint and multinational operations, and (c) emerging DoD initiatives to ensure officers possess foreign language proficiency and knowledge of cultural customs, traditions, and history.

Imperatives and Implications Impacting Selection System Content. By virtue of their influence on selection system strategy, the forces and challenges facing the Air Force can be expected to exert an influence on what (and how many) attributes to assess when selecting junior officers. As other military organizations have discovered (e.g., for the Army – see Ford et al., 2000; Knapp, McCloy, & Heffner, 2004; Sager et al., 2004) and as evident from Table 7, the expanded and increased complexity of requirements for officer success implicates a broader range and greater number of attributes than are currently assessed. Thus, maximizing the prediction of mission critical outcomes, and thereby the Air Force's ROI, is likely to require changes in the number and kinds of attributes assessed.

Imperatives and Implications Impacting Selection System Methods, Design, and Implementation. Finally, and consistent with the preceding sections, there are indications that the forces the Air Force faces could influence how it assesses officer candidates and other design choices. Table 8 summarizes the imperatives flowing from these forces and their potential implications. For example, greater diversity in officer candidates' educational experiences (e.g., private versus public versus home-schooled) is likely to impact the efficacy of using GPA, as such diversity will render the ability to make valid, cross-candidate comparisons more difficult.

Table 5. Comparison of Air Force Enduring Leadership Competencies (Tactical-Operational Levels) to Other Services

Air Force Enduring Leadership Competencies		**Service**		
		Army [1]	*Navy* [2]	*Marine Corps* [2]
Leading the Institution	Embrace Change and Transformation		✓	
	Drive Execution	✓	✓	
Leading People/Teams	Drive Performance Through Shared Vision, Values, and Accountability	✓	✓	✓
	Influence Through Win/Win Solutions	✓	✓	✓
	Mentor and Coach for Growth and Success	✓	✓	✓
	Promote Collaboration and Teamwork	✓	✓	✓
	Partner to Maximize Results	✓	✓	
Personal Leadership	Exercise Sound Judgment	✓	✓	✓
	Adapt and Perform Under Pressure	✓	✓	✓
	Inspire Trust	✓	✓	✓
	Lead Courageously	✓	✓	✓
	Assess Self	✓	✓	✓
	Foster Effective Communication	✓	✓	✓

Notes. [1] From FM22-100, *Army Leadership* (U.S. Department of the Army, 1999); cited in Horey & Fallesen (2003).
[2] Cited in Horey & Fallesen (2003).
Because they are not relevant to the performance of junior officers, several of the specific competencies (e.g., ―Shape Air Force Strategy and Direction”) within the ―Leading the Institution” meta-competency have been excluded from table.

Table 6. Performance Imperatives and Representative Competencies Identified by Future-Oriented Research on 21st Century Military Leadership Requirements

Performance Imperative	**Representative Competencies**	**Rationale(s)**
Cognitive	• Decision-Making • Creativity & Innovation • Problem-Solving	Growing complexity and knowledge intensiveness of leader's own work (and that of others); organization- and Service-wide efforts to promote (and sustain) a culture of creativity and innovation; increased speed and operational tempo requiring leaders to make decisions at consistently faster rates.
Informational	• Information & Knowledge Management • Oral Communication • Written Communication	New and evolving command and control structures necessitating information and monitoring sharing across different levels, functional areas and commands; increased volume and complexity of information; effective communication increasingly critical to ensure shared understanding of tasks, the situation, and so on.
Social	• Cultural Adaptability-Competence • Developing & Coaching Others/ Team Building • Direction Setting • Diversity & Conflict Management • Influencing & Motivating Others • Operational Execution-Management • Role-Modeling • Selfless Service Orientation/ Organizational Citizenship • Teamwork	Increased need to work in a team-based environment and collaborate effectively with others to achieve mission success; greater demand to lead and work in multicultural, multiracial and mixed-gender teams; growing responsibility for managing and leveraging diversity; increased requirement for collective leadership (at all levels).
Personal	• Adaptability • Commitment to High Performance • Initiative • Integrity • Safety Consciousness • Self-Directed (or Continuous) Learning • Self-Management	Increased speed and operational tempo; greater task variety and higher performance demands; strong requirement for continuous learning and the need to independently manage one's (personal and professional) development; growing emphasis on individual's role in balancing and managing his/her own personal matters and well-being.
Political	• Boundary Spanning-Management • Building Coalitions & Partnerships • Extending Influence • Leading & Implementing Change • Negotiating & Persuading Others	Shrinking of traditional boundaries due to technology and changing organizational structures; increased need to lead, collaborate, and partner with extra-organizational (and non-military) components to get things done; greater responsibility for implementing and facilitating organization- and Service-wide transformation efforts.
Technological	• Computer-Technological Competence	Increased use of computers and other technologies to complete work (individually and collectively).
Financial	None	N/A

Table 7. Overview of Attributes Related to Air Force Leader Competencies

Category	Attributes	Representative Air Force Competencies
Aptitudes and Abilities (Cognitive, Psychomotor, Sensory, Spatial)	• Abstract-Analytical Reasoning • Creative-Divergent Thinking • Critical Thinking • Cognitive-Integrative Complexity • General Mental Ability (GMA) • Information Processing & Analysis • Moral Reasoning • Practical Intelligence • Problem-Solving • Social-Emotional Intelligence • Systems Thinking & Planning	• Leading the Institution • Leading People/Teams • Personal Leadership (Exercise Sound Judgment; Inspire Trust; Foster Effective Communication)
Cross-Functional Skills	• Basic Computer Skills • Conflict Management & Resolution Skills • Decision-Making Skills • Dynamic Information Processing & Analysis Skills • Metacognitive & Motivational Skills • Negotiation Skills • Persuasion Skills • Problem-Solving Skills • Self-Directed Learning & Development Skills • Self-Management Skills • Situational Awareness Skills • Social-Interpersonal Skills • Tacit Knowledge • Teamwork Skills • Verbal-Written Communication Skills	• Leading the Institution • Leading People/Teams • Personal Leadership (Exercise Sound Judgment; Adapt & Perform Under Pressure; Inspire Trust; Assess Self; Foster Effective Communication)
Personal Qualities (Personality Traits, Motives, Values, Interests)	• Achievement Motivation • Adaptability • Affiliation Motivation • Agreeableness • Conscientiousness • Creativity • Cultural Tolerance • Dependability • Decisiveness • Dominance-Surgency • Energy Level-Potency • Extraversion • Emotional Maturity-Stability • General Self-Esteem • Generalized Self-Efficacy • Initiative-Proactive Personality • Intellectance • Integrity • (Internal) Locus of Control • Motivation to Lead-Socialized Power Motive • Motivation to Learn • Openness to Experience • Persistence/Self-Reliance • Risk Propensity • Self-Discipline • Self-Monitoring/Social Perceptiveness • Service Orientation • Sociability • Stress Tolerance • Team Orientation • Tolerance for Ambiguity	• Leading the Institution • Leading People/Teams • Personal Leadership (Exercise Sound Judgment; Adapt & Perform Under Pressure; Inspire Trust; Lead Courageously; Assess Self)

Table 8. Summary of the Performance Imperatives and Implications Impacting Selection System Methods, Design, and Implementation

Performance Imperative	Rationale	Implications
Greater demand to demonstrate the value-added/return-on-investment (ROI) from using specific selection system components.	Increased centrality and strategic importance of officer selection to the Air Force competing (and potentially conflicting) with current and continued budgetary pressures.	• Employ selection methods that maximize value/ROI relative to potential alternatives.
Growing need to facilitate the achievement of Total Force goals.	Total Force transformation requiring cross-component and joint service selection systems and tools.	• Design selection system (and its components) to ensure a sufficiently diverse and sizeable cohort of officer candidates to meet Force mix requirements. • Employ selection methods that can be deployed across Air Force components and the sister services (Army, Navy, Marine Corps).
Need to facilitate the accomplishment of recruitment goals.	Increased pressure on attracting and recruiting junior officer candidates.	• Design selection system (and its components) for applicant —pre-selection” or —self-deselection.” • Employ selection methods (and other system design and implementation features) that promote favorable reactions (e.g., perceived fairness) and acceptance among officer candidates and prospective applicants.
Need to link selection and training.	Increased potential for misalignment (or gaps) between prospective candidates' attributes/ qualifications and officer (and/or technical specialty) performance requirements.	• Employ selection methods (and other system design and implementation features) that facilitate future training and development (e.g., providing diagnostic feedback to prospective candidates and trainers).
Need for stakeholder "buy in" and support.	Essential to the successful implementation and sustainment of selection system (and its components).	• Design selection system (and its components) to be efficient to deploy and implement with reasonable investments of time and resources. • Employ selection methods that represent a —good” cultural fit and can meet the commonsense criteria (e.g., perceived as Air Force- and officer-relevant) of Air Force stakeholders.
Dealing with a more diverse officer candidate pool.	Greater diversity in youth population and the kinds of qualifications prospective candidates bring to the Air Force (e.g., educational experiences).	• Employ selection methods that minimize adverse impact without sacrificing quality (where feasible). • Employ standardized selection methods.
Facilitates, rather than delays, future changes to officer selection resulting from evolving force needs.	Evolving and future force needs (e.g., from operational —lessons learned”) may require periodic changes.	• Design selection system (and its components) to be sufficiently flexible and responsive to changing and evolving force needs and corresponding performance requirements.
Increasingly visible need to make effective use of technology to select junior officers.	Proliferation and increased use of technology by competitors in their selection processes, coupled with Air Force's need to promote a desirable organizational image (e.g., high tech force).	• Design selection system (and its components) to leverage existing and emerging technology to enhance officer selection (where feasible).

What Makes an Air Force Officer Selection System Successful? A Review and Integration

Overall, our review of the forces and challenges affecting the Air Force as it transitions to a 21st Century force indicates they carry several important implications for how the Air Force selects junior officers. Further, many of these forces are exerting comparable effects on the other Services (cf. Ford et al., 2000; Horey et al., 2004; Sager et al., 2004) and non-military organizations (cf. Higgs et al., 2000; Ilgen & Pulakos, 1999; Pearlman, 2003; Pearlman & Barney, 2000). As an integrative review of the preceding sections, Table 9 summarizes the main performance imperatives facing the Air Force and its goals for achieving officer selection system success. These goals became the effectiveness criteria against which we evaluated the current selection systems (and their components) in place at AFROTC and OTS.

Step 2: Identifying Critical Gaps in Current Officer Selection Systems

Having identified the effectiveness criteria important to promoting selection system success, we evaluated the current officer selection systems (and their components, such as the AFOQT) in use at AFROTC and OTS. This evaluation was based on a comprehensive review of relevant (a) Air Force regulations, instructions, and directives (e.g., AFI 36-2005; AFOATSI 36-2013; AFRS Procedural Guidance Messages), (b) application forms (e.g., AF IMT 56), and (c) technical reports, conference papers, and publications (e.g., ANSER, 2005; Arth, 1986; Cowan, Barrett, & Capt, 1989; Cowan, Barrett, & Wegner, 1990; Hartke & Short, 1988; Shore, Gould, Ree, Alley, & Skinner, 2003). In addition, we consulted with AFROTC and AFRS representatives with extensive knowledge of their respective selection platforms. What follows are the most critical gaps identified from this review, organized by selection system (a) strategy, (b) content, and (c) methods, design, and implementation.

Selection System Strategy

Overall, our review identified several critical gaps (or potential gaps). First, evidence is limited that the current selection systems (and their components) are meaningfully and measurably linked to current and emerging Air Force competency (performance) requirements, specifically those embodied in Air Force doctrine (AFDD 1-1). Generally speaking, the selection systems (both AFROTC and OTS) currently in place for non-rated, line officers have largely remained unchanged for over 20 years. Further, and consistent with the system in place at the USAF Academy (see ANSER, 2005), there is no formal process at either AFROTC or OTS for clearly and meaningfully specifying the linkages between the predictors and assessment measures in use to official competency (performance) requirements. As a result, how effectively commissioning is generating the kinds of officers the Air Force needs is not clear, since efforts to regularly evaluate selection system success – beyond meeting production goals (or end strength requirements) – are generally limited. While AFOATS conducts an annual survey of recent graduates and their immediate supervisors, which includes items assessing the graduate's professionalism and leadership, the survey potentially suffers from many of the same limitations as Officer Performance Reports (OPRs) (e.g., significant range restriction, uniformly and highly positive ratings, and exclusive focus on individual-level effectiveness). For the Air Force, ensuring there is a strong and measurable link between its HR systems (officer selection, training, and so on) and its goals and performance imperatives is important to ensure further

Table 9. Performance Imperatives and Goals Driving the Success of Air Force Officer Selection

Performance Imperatives	**Air Force Goals**
Value-Added/ *Return on Investment (ROI)*	• The selection system (and its components) maximizes ROI by optimally balancing officer quality with costs (i.e., development, operational/administrative, etc.).
Force Quality & Impact	• The selection system (and its components): o are systematically linked to enduring leadership competencies, core Air Force values, and projected capabilities to ensure the selection of officers with the qualifications needed to meet current and emerging Force needs; o exerts and can demonstrate a measurable impact on critical individual, team (or unit), and organizational outcomes.
Force Diversity & Manpower	• The selection system (and its components): o ensures a sufficiently diverse and sizeable cohort of quality officers to flexibly meet changing Force mix and manpower requirements; o can be employed across Air Force components (Active, Reserve, Civilian) and the other Services (Army, Navy, Marine Corps) to achieve Total Force mix requirements.
Implementation & Management	• The selection system (and its components) are efficient to use and manage by the intended stakeholders (e.g., AFROTC, AFRS, OTS, and so on) with reasonable investments of time and resources.
Integration & Standardization	• The selection system (and its components): o can be programmatically aligned with other HR systems (e.g., recruiting) to support and leverage Air Force investments; o can be implemented and sustained across officer commissioning sources (AFROTC, OTS, and USAFA) to facilitate and promote standardization—as needed— within the Air Force and across the sister services (Army, Navy, Marine Corps).
Flexibility & Responsiveness	• The selection system (and its components) are sufficiently flexible and responsive to changing Force needs (e.g., from operational ―lessons learned”) so as to be ready when (and as) needed to facilitate rather than delay needed changes in officer selection.
Stakeholder “Buy In”	• The selection system (and its components): o meets commonsense criteria (i.e., perceived as Air Force- and officer-relevant, easy-to-use, etc.) and are positively received by internal Air Force stakeholders; o represents a good cultural ―fit” with Air Force values; o promotes favorable reactions (i.e., perceived as fair, strongly predictive of officer performance, etc.) among prospective officer candidates; o engenders positive image to ensure the attractiveness of Air Force to a changing and increasingly diverse officer candidate pool.
Technology	• The selection system (and its components) leverages existing and emerging technology to facilitate, but *not* drive, Air Force needs and goals (above).

success (cf. Higgs et al., 2000; Wright & Boswell, 2002). While none of this should be interpreted to mean that the current systems have not been serving the Air Force well, it does suggest the potential for serious gaps that could adversely impact its continued ability to select the right people with the right qualifications to be leaders.

Second, and related to the first point, there have historically been and continue to be substantive and practical differences in how AFROTC and OTS define and assess officer quality for purposes of selection. Although realigned under AFOATS (in 1997) to facilitate and ensure coordination in the commissioning of junior officers, the selection systems currently in place at AFROTC and OTS are essentially the same as they were prior to realignment. Chief among these differences is the relative emphasis AFROTC and OTS place on selecting generalists (i.e., officership) versus specialists (i.e., technical expertise). Specifically, whereas AFROTC places greater emphasis on Officership, OTS more strongly takes into consideration prospective candidates' technical (or functional) expertise – as exemplified by its practice of dividing selection boards into sub-panels organized by function (e.g., Critical Technical, Technical, and Non-Technical) and in some cases career field. While these differences are not inherently counterproductive, how and to what degree AFROTC and OTS *should* differ is an open and important question since it carries significant implications for the success of Force Shaping and similar initiatives, where effectively managing the mix of officers produced, and in what numbers, is critical. Ultimately, the answer to this question will depend on each commissioning source's role in meeting Force needs (both historically and in the projected future), the nature and kinds of populations they serve, how officer selection intersects with other Force goals (e.g., recruitment and visibility), and so on. Consistent with the preceding point, coming to a solution will require the careful specification of what the selection systems should be predicting – and how best to align officer selection systems with these specifications – so that the pros and cons of differences across commissioning sources can be effectively weighted.

Third, promoting Force diversity has been and will continue to be challenging until there are sufficiently clear specifications and comprehensive operationalizations of diversity, such that officer selection systems can be structured and aligned accordingly. Over the next several decades, these challenges are expected to increase in importance as changing demographic and economic factors create increased competition for leader talent, especially in select occupations (cf. Brazell & Sharon, 2004; Census, 2004; Galway et al., 2005; Hosek, Mattock, Fair, Sharp, & Totten, 2004; Michaels, Handfield-Jones, & Axelrod, 2001). As evident from Appendix A, recent Air Force efforts to access and commission a demographically (i.e., gender, race/ethnicity) diverse officer pool compare favorably, on average, to the other Services. Nevertheless, the success of these efforts has varied by commissioning source. In addition, in several cases, sizeable gaps between Air Force and the civilian population as a whole persist. Consistent with the preceding discussion, comparable results have been found when diversity is considered in terms of airmen capabilities (e.g., competencies and knowledge, skills, and abilities, KSAs) (e.g., Galway et al., 2005). Since there is limited evidence to suggest that the assessments in use for selecting officers are systematically biased (e.g., Carretta & Ree, 1998; Roberts & Skinner, 1996), achieving and sustaining Force diversity will require (a) a sufficient and comprehensive specification of diversity, (b) identification of the right diversity metrics, (c) models of the linkages between diversity metrics and performance-relevant outcomes, and (d) aligning officer selection systems to facilitate and support Force diversity initiatives (e.g., Ployhart, 2004; Ployhart & Schneider, 2002; Sacco & Schmitt, 2005).

Finally, there is limited evidence that the current selection systems are sufficiently integrated and aligned with other HR systems – specifically officer recruitment and training – to substantively contribute to and inform the processes associated with these systems. Because recruitment and training are more costly than selection (Pearlman & Barney, 2000), not taking advantage of the selection process to maximize the Air Force's collective ROI in these different systems represents a missed opportunity. For example, since the early 1990s, attracting applicants to military service has grown more difficult and costly (Bachman, Segal, Freedman-Doan, & O'Malley, 2000; GAO, 2005; Kilburn & Asch, 2003; Orvis & Asch, 2001; Sackett & Mavor, 2003). One potentially contributing factor is that more and more prospective applicants – particularly high quality applicants – do not view the Services as a place where they can acquire transferable, job-related skills needed for a successful, financially viable career outside of the military (Kilburn & Asch, 2003; Sackett & Mavor, 2003). While comparatively speaking, the Air Force tends to be perceived more favorably than the other Services (cf. Department of Defense, 2005), how current officer selection systems – AFROTC and OTS included – contribute to these attitudes and perceptions is not well understood. More importantly, how these selection systems could best be designed and implemented to sustain and enhance the attractiveness of military service and of becoming an Air Force officer have similarly received limited study. As mentioned previously, these issues are expected to grow in criticality as emerging demographic and economic trends foster increased competition for leader talent, especially in select occupations (Brazell & Sharon, 2004; Galway et al., 2005; Hosek et al., 2004; Michaels et al., 2001). Nor are these considerations limited to recruitment. For instance, at present none of the assessments employed by AFROTC or OTS, such as the AFOQT, is used to formally and programmatically deliver meaningful feedback to officer candidates. Doing so would be advantageous because it would (a) assist in the ―selecting out" of less qualified candidates, and/or (b) better prepare those candidates selected for subsequent training and development (e.g., by providing candidates with a targeted diagnosis of their strengths and weaknesses with recommendations on relevant developmental experiences).[13] Consequently, despite Air Force expense in assessing officer candidates, it generally does not appear that this investment is sufficiently leveraged to support a broader range of personnel decisions (e.g., training, classification) beyond selection.

Content

Several critical gaps (or potential gaps) were identified from our review. First, in several cases, precisely what attributes the current selection systems are measuring is not always clear – excluding the AFOQT and other standardized tests (SAT/ACT) of general and specific mental abilities, whose construct validity is well-established (e.g., Carretta & Ree, 1996; Carretta, Retzlaff, Callister, & King, 1997; Diehl, 1986; Earles & Ree, 1991; Sperl, Ree, & Steuck, 1992). Generally speaking, a number of the attributes assessed by the current selection systems are essentially collections of specific measures whose relationship to an underlying attribute(s) predictive of future officer performance (or retention) has not been clearly specified or established. This situation is problematic because, all other things being equal, properly specified and well-defined predictor attributes can greatly enhance the selection and/or construction of measures that optimize an organization's ROI (Guion, 1998). For example, the Relative Standing Score (RSS) employed by AFROTC is presumably meant to assess ―officer potential." Based on previous research, RSS appears to be potentially measuring something, since (a) it is not strongly correlated with other AFROTC-PSP predictors, such as AFOQT-AA, SAT/ACT scores, or GPA (*r*s = .10s-.30); and (b) relative to these other predictors (shown in parentheses are criterion-related validity for AFOQT-AA versus validity for RSS), it is differentially related to POC-related criteria (*r*s = .10s vs. .10-.20s) (e.g., student performance rating, course completion) and technical training performance criteria (*r*s = .30s vs. .00) (Cowan et al., 1989). Nevertheless, what precisely RSS is measuring remains unclear, as its substantive content has not been explicitly articulated and operationalized (see AFOATSI 26-2013, p. 4). This absence of a clear and consistent conceptual definition could be introducing error and possibly other contaminants into RSS scores, thereby explaining why empirical estimates of its predictive validity do not necessarily match its prescribed weight (e.g., Cowan et al., 1989). Given the practically and statistically significant role ―officer potential" (RSS) plays in AFROTC officer selection, addressing this issue is important. We point this out not to completely rule out ―officer potential" (or RSS) as a predictor – indeed RSS may very well be measuring ―officer potential" and one or more of the attributes previously identified (see Table 7). However, we suspect its value will likely continue to be undercut until substantive, and performance relevant, specifications of what ―officer potential" is are generated to ensure standardized and construct valid assessment.

---

[13] For instance, Carretta & Ree (1995) observed that AFOQT subtests were differentially predictive of success at different points in pilot training.

Second, as evident from juxtaposing Tables 2 and 7, there are several alternative predictor attributes linked to one or more of the Air Force competency (performance) requirements for officers not captured by the current selection platforms – or related to the preceding point, not assessed in a (construct) valid way. Overall, and as summarized in Tables 10 and 11, while the predictors currently comprising these systems, specifically aptitude-ability, do a reasonably good job of predicting commissioning program success (e.g., grades, instructor ratings, and so on) (*r*s = .10-.30s), they are less efficacious in predicting post-commissioning effectiveness (*r*s $\leq$ .15).[14] This pattern is consistent with other organizational research, where general mental ability (GMA) operationally explains, on average, about 4% of the variance in who becomes a leader and how effectively they perform in an actual leadership role (Judge et al., 2004). Similarly, when aggregated into an order of merit score, the predictors used by AFROTC and OTS for selecting officer candidates can operationally explain roughly 1% to 27% (mostly for OTS) depending on the performance criterion (i.e., emergence or post-commissioning effectiveness).[15] While this level of prediction is informative, these findings indicate that there are potentially other attributes that could add to the prediction of *leader* (officer) potential and could possibly do so more efficiently.[16] For example, and as illustrated in Tables 10 and 11, personality attributes (e.g., the Big Five) can produce, on average, operational validities (for leader emergence and effectiveness performance criteria) that rival, or are nearly *double*, those of many of the predictors currently in use. While research on other alternative predictor attributes, such as Cognitive-Integrative Complexity, is less extensive, initial research with Army officers is similarly promising (cf., Lewis, 1995; McIntyre, Jordan, Mergen, Hamill, & Jacobs, 1993).

---

[14] Beyond statistical artifacts, one viable explanation for these differences is that the knowledge, skills, and abilities (KSAs) acquired in commissioning mediate the influence of aptitude-ability on on-the-job effectiveness (e.g., Borman, White, Pulakos, & Oppler, 1991; Borman, Hanson, Oppler, Pulakos, & White, 1993). We were unable to test such a causal model, however, since correlations among emergence and effectiveness criteria were not reported (Cowan et al., 1989; Cowan et al., 1990)

[15] Taking into consideration various statistical artifacts (specifically criterion unreliability and range restriction), which typically can increment operational validities by about 50%, one would still be looking at explaining roughly 2% to 40% (commissioning success) of the variance in performance criteria.

When considering the complexity of the assignment officer selection is tasked with, the necessity of expanding the predictor space makes sense. While their relative importance varies across commissioning sources, junior officer selection is intended to inform multiple selection decisions, including who (a) will successfully complete his/her academic degree program, (b) will successfully complete his/her commissioning program, whose content and performance requirements differ from that of their degree program, (c) possesses the requisite attributes to perform effectively as a junior officer, coupled with the development potential to grow into an effective *senior* officer, and (d) represents a reasonably good ―fit" with Air Force culture and its values. Collectively, these decisions implicate a multidimensional, and potentially conflicting, set of outcomes that are not highly correlated, and whose predictors will vary (Borman, Hedge, Ferstl, Kaufman, Farmer, & Bearden, 2003; Guion, 1998; Ilgen & Pulakos, 1999; Judge & Piccolo, 2004; Murphy & Shiarella, 1997; Pearlman, 2003; Pearlman & Barney, 2000). Therefore, maximizing prediction will require an expanded and more diverse set of predictors, as research in both military and non-military organizations illustrates (cf. Campbell & Knapp, 2001; Geraghty & Collins, 2003; Knapp et al., 2004; Motowidlo & Van Scotter, 1994; Pulakos, Schmitt, Dorsey, Arad, Hedge, & Borman, 2002; Van Scotter & Motowidlo, 1996).

---

[16] For instance, OTS selection boards have up to 50+ different indicators (e.g., AFOQT scores, college transcripts, letters of recommendation, recruiter/interviewer ratings, and so on) to consider when evaluating prospective candidates – even though the operational validities of several of these (e.g., letters of recommendation), relative to possible alternatives (e.g., personality), is substantially less.

Table 10. Comparison of Operational Criterion-Related Validities for AFROTC Predictors to SDI and Meta-Analytic Findings from Organizational Research

	Emergence			Effectiveness			
Predictor	Completion	DG	Rating	OER	Overall Leadership	Potential	Motivation
AFROTC							
Aptitude-Ability							
AFOQT-AA	.06	.15	.12	.09	.07	.14	.05
AFOQT-Q	.07	.12	.08	.11	.05	.12	.03
AFOQT-V	.04	.13	.13	.04	.03	.06	.01
SAT	.07	.16	.12	.09	.06	.13	.04
Academic Performance-Experience							
Cumulative GPA	.17	.31	.24	.05	.07	.06	.04
AFROTC GPA	.10	.19	.18	.06	.06	.10	.07
GMC Credit	.03	.04	.05	.06	.06	.06	.05
Officer Potential							
Unit Commander's Rating	.11	.21	.19	.06	.10	.11	.11
Unit Commander Ranking (1-50)	.11	.22	.19	-.01	-.03	-.02	-.02
Unit Commander Ranking (All)	.01	.01	-.01	.05	.07	.10	.10
Order of Merit (OM)	.14	.28	.22	.10	.10	.14	.06
Self-Description Inventory (SDI)							
Agreeableness		--		--	.40	.31	.39
Conscientiousness		--		--	.23	.17	.11
Emotional Stability		--		--	.34	.33	.24
Extraversion		--		--	.13	.20	.17
Openness to Experience		--		--	.27	.29	.20
Findings from Comparative Organizational Research							
General Mental Ability (GMA)		.16			.14		
Personality							
Agreeableness		.03			.14		
Conscientiousness		.23			.11		
Emotional Stability		.17			.16		
Extraversion		.24			.17		
Openness to Experience		.17			.17		

Note. All validities are uncorrected for unreliability (predictor and criterion) and range restriction. Validities for AFROTC predictors from Cowan et al. (1989). Validities for comparative organizational research are sample-size weighted mean observed validities reported by Judge et al. (2002) and Judge et al. (2004). SDI validities from Christal et al. (1997, p. 8), collected from sample of junior officers who attended AFROTC or SOS (*n* = 440). DG = Distinguished Graduate. OER = Officer Effectiveness Reports. To facilitate interpretation, signs (+/-) of OER validities were switched from those originally reported, so that more positive ratings are associated with higher predictor scores. *Overall Performance*, *Potential (for Advancement)*, and *Motivation (to Perform)* were experimental appraisal forms developed specifically for use in Cowan et al. (1989, 1990).

Table 11. Comparison of Operational Criterion-Related Validities for OTS Predictors to SDI and Meta-Analytic Findings from Organizational Research

	Emergence			**Effectiveness**			
Predictor	Completion	Grade/DG	Rating	OER	Overall Leadership	Potential	Motivation
OTS							
Education-Aptitude							
AFOQT-AA	.09	.36(.05)	.08	-.03	.04	-.02	-.05
AFOQT-Q	.04	.15(.01)	-.02	-.02	.00	-.06	-.09
AFOQT-V	.12	.41(.07)	.15	-.02	.06	.02	-.01
GPA	.03	.30(.21)	.29	.00	-.01	.13	.10
Experience							
Work Experience-Nonmanagerial	.05	.04(.05)	.12	.02	--	--	--
Work Experience-Managerial	.01	.00(.02)	.02	-.01	--	--	--
Nonmilitary Awards	.00	.03(.08)	.06	.00	--	--	--
Nonmilitary Achievements	.00	-.01(.03)	.00	-.04	--	--	--
Extracurricular Activities	.01	-.03(.01)	.00	-.05	--	--	--
Potential/Adaptability							
Recruiter Evaluation-Overall	.00	.03(.04)	.06	.04	.06	.02	.03
Recruiter Evaluation-Communication Skill	-.01	-.04(.05)	.03	-.02	--	--	--
Recommendation Letter-Civilian	-.10	-.06(-.04)	-.11	-.02	--	--	--
Recommendation Letter-Military	.08	.00(.04)	.10	.01	--	--	--
Self-Description Inventory (SDI)							
Agreeableness	--			--	.40	.31	.39
Conscientiousness	--			--	.23	.17	.11
Emotional Stability	--			--	.34	.33	.24
Extraversion	--			--	.13	.20	.17
Openness to Experience	--			--	.27	.29	.20
Findings from Comparative Organizational Research							
General Mental Ability (GMA)	.16			.14			
Personality							
Agreeableness	.03			.14			
Conscientiousness	.23			.11			
Emotional Stability	.17			.16			
Extraversion	.24			.17			
Openness to Experience	.17			.17			

Note. All validities are uncorrected for unreliability (predictor and criterion) and range restriction. Validities for OTS predictors from Cowan et al. (1990). Validities for comparative organizational research are sample-size weighted mean observed validities reported by Judge et al. (2002) and Judge et al. (2004). DG = Distinguished Graduate. OER = Officer Effectiveness Reports. To facilitate interpretation, signs (+/-) of OER validities were switched from those originally reported, so that more positive ratings are associated with higher predictor scores. *Overall Performance*, *Potential (for Advancement)*, and *Motivation (to Perform)* were experimental appraisal forms developed specifically for use in Cowan et al. (1989, 1990).

Methods, Design, and Implementation

From our review, several critical gaps (or potential gaps) were identified. First, and as pointed out by others (e.g., Cowan et al., 1989, 1990), a number of the officer performance measures or metrics (e.g., OPRs/OERs) used for purposes of either validating and/or evaluating the success of the current selection systems (and their components) are potentially problematic. Specifically, a number of these measures (a) are not explicitly and/or measurably linked to competency (performance) requirements, (b) are potentially deficient, and (c) do not sufficiently differentiate among officers. To illustrate the practical impact of these issues consider the observed differences in the operational validities between the AFOQT and SDI reported in Tables 10 and 11. While these differences likely reflect actual, substantive differences in predictive validity, they also reflect differences in the quality of the performance measures used. Whereas the SDI validities were estimated using sound, well-constructed performance measures developed as part of the Air Force's Job Performance Measurement System (JPMS) project (Hedge & Teachout, 1986), AFOQT validities were based on OERs.[17] As evident from these tables, conclusions about the comparative ROI of alternative predictors can be greatly influenced by the nature and quality of the performance measures used. Equally as important, most of the performance measures currently in use by Air Force focus exclusively on the individual level-of-analysis. Excluding outcomes at higher levels could preclude important information about selection system success, since Air Force culture and emerging competency (performance) requirements emphasize that junior officers are increasingly expected to contribute to team- and unit-level outcomes. Further, many of the Air Force's human capital challenges – and thereby its most critical decisions now and in the future – depend on being able to effectively and measurably link lower-level processes (e.g., officer selection) to higher-level outcomes. For example, diversity – whether based on Airmen demography and/or capabilities – is inherently a higher-level phenomenon (Ployhart, 2004). Therefore, how best to manage a cohort (or pool) of officers can only be understood in the aggregate. In brief, it requires measures targeting the pool's composition (e.g., percentage of officers within a cohort with a desired set of competencies and KSAs) and relating said measures to organizationally-relevant outcomes. Because of this, how best to select a *cohort* of officers with (some) desired composition of attributes could differ substantially from how best to select *individual* officers (e.g., Jones, Stevens, & Fischer, 2000). Understanding these issues carries substantial implications for how best to design and implement an officer selection system.

Second, both AFROTC and OTS currently employ selection models that, as presently designed, essentially force the Air Force to make tradeoffs that in the long-term are neither desirable for the prospective officer candidate or the Air Force. Specifically, both AFROTC and OTS follow a compensatory model in which officership and technical expertise are considered simultaneously. Consequently, exceptionally high technical expertise can compensate for poor officer potential and vice versa. As documented by others (cf. Carretta, 2000; Weeks, 2000), the result is the selection of candidates with insufficient aptitude and/or skills needed to successfully complete technical training – and similarly, highly technically proficient candidates lacking the requisite attributes to be successful officers. Ultimately, neither outcome is productive for the Air Force; both are costly in terms of time and resources. More importantly, such tradeoffs (in the aggregate) could substantially impact the Air Force's ability to meet its talent needs and Force mix requirements. How best to resolve this issue carries significant implications for a range of critical design choices, such as how the Air Force weights different predictors when selecting officer candidates, under what model (e.g., compensatory, multiple hurdle), and so on.

---

[17] Similarly, compared to the OERs, the experimental performance measures developed and employed by Cowan et al. (1989, 1990) demonstrated greater coverage of leader (officer) performance requirements and a stronger ability to differentiate among officers in terms of their performance.

Third, with the exception of the AFOQT and similar standardized assessments, AFROTC and OTS employ several unstandardized (or unstructured) assessments, where the absence of substantive and clear specifications of what they intend to measure are potentially undercutting their ROI – specifically, RSS for AFROTC and recruiter/interviewer evaluations for OTS. Similarly, OTS continues to rely on letters of recommendation, even though research shows that such recommendations are rarely negative and therefore not especially informative when evaluating applicants (Aamodt & Williams, 2005). This point has at least partially been borne out by existing Air Force research (e.g., Cowan et al., 1990), where civilian letters of recommendation are negatively related to OTS success. This situation is problematic because these assessments, particularly RSS, (a) factor into and contribute significantly to who gets selected (e.g., Cowan et al., 1989, 1990; Weeks, 2000), and (b) are intended to provide assessments of attributes not measured by the AFOQT and similar academically or cognitively oriented predictors (e.g., GPA). In the case of RSS, for example, what —officer potential" means could vary considerably across unit commanders. Because RSS scores are essentially normed within units, this raises the potential for significant differences in what the same RSS score reflects between different units. While such unstructured assessments might first appear cost-effective, we suspect they are not since they probably require considerable time and resources to administer, score, and/or evaluate. Addressing these issues is important because they could potentially yield reasonable dividends at minimal effort and expense to the Air Force, relative to more expensive alternatives.

Air Force Officer Qualifying Test (AFOQT)

Since 1951, the AFOQT has primarily been used by the Air Force to (a) select officer candidates (rated and non-rated; line and non-line) for AFROTC and OTS, and (b) classify rated candidates for undergraduate pilot and navigator (combat system operator) training. Operational since June 2005, the current version of the AFOQT (Form S) consists of 11 cognitive subtests; two of which – Hidden Figures (HF) and Rotated Blocks (RB) – are used exclusively for research purposes (Shore et al., 2003; Weissmuller et al., 2004). In addition, Form S includes an experimental test – the Self-Description Inventory (SDI+) – a self-report personality measure assessing the Big Five domains and facets, plus two compound traits reflecting *Service Orientation* and *Team Orientation*. Table 12 summarizes the current content and structure of Form S and how the different subtests map to the five AFOQT composites employed by AFROTC and OTS for officer selection and categorization.

Table 12. Summary of AFOQT Form S Content

Subtest	Description	Number of Items	Testing Time (in Minutes)	AFOQT Composites				
				Pilot	Navigator	Academic Aptitude	Verbal	Quantitative
Verbal Analogies (VA)	Assesses ability to reason and recognize relations between words.	25	9		X	X	X	
Arithmetic Reasoning (AR)	Measures understanding of arithmetic relations expressed as word problems.	25	30	X	X	X		X
Word Knowledge (WK)	Measures ability to understand written knowledge through use of synonyms.	25	6			X	X	
Math Knowledge (MK)	Measures knowledge of mathematical terms, formulas, and relations.	25	23	X	X	X		X
Instrument Comprehension (IC)	Assesses ability to ascertain aircraft attitude from illustrations of flight instruments.	20	9	X				
Block Counting (BC)	Assesses spatial ability through analysis of three-dimensional representations of a set of blocks.	20	5		X			
Table Reading (TR)	Assesses the ability to quickly and accurately extract information from tables.	40	9	X	X			
Aviation Information (AI)	Measures knowledge of general aviation terminology and concepts.	20	9	X				
General Science (GS)	Assesses knowledge and understanding of scientific terms, concepts, principles, and instruments.	20	11		X			
Rotated Blocks (RB)	Measures spatial aptitude by requiring mental manipulation and rotation of objects.	15	15					
Hidden Figures (HF)	Measures spatial ability by requiring the discovery of simple figures embedded in complex drawings.	15	10					
Self-Description Inventory (SDI+)	Assesses standing on Big Five personality dimensions (and their facets), plus Service and Team Orientation.	220	40					
Totals		**470**	**213**					

After reviewing the AFOQT's current and proposed content and relevant research (e.g., Christal, Barucky, Driskill, & Collis, 1997; Gould et al., 2003), we identified the following gaps (or potential gaps) as most critical.

First, while the inclusion of the SDI+ represents a welcomed and needed effort to substantively expand the number and kinds of predictor attributes assessed by the AFOQT, the efficacy of continuing to require all prospective candidates to expend limited test-taking time completing tests (e.g., IC, BC, TR, AI, and GS) used exclusively in the selection and categorization of rated officers (pilots, navigators) is not clear. As can be seen from Table 12, for non-rated officer candidates, of the nine cognitively oriented and operational subtests currently comprising the AFOQT, five (about 40 minutes worth) expressly target aptitudes and skills exclusive to flying specialties. The remaining four provide an assessment of GMA (cf. Carretta & Ree, 1996; Earles & Ree, 1991). For various practical reasons, greatly expanding the existing AFOQT beyond a half-day timeframe is understandably not feasible. Since the greatest practical increments in predictive efficacy are expected to accrue from measuring predictor attributes beyond GMA (cf. Borman et al., 2003; Guion, 1998; Pearlman & Barney, 2000; Schmidt & Hunter, 1998), and provided that the ultimate goal of the AFOQT is to serve as a standardized, centralized assessment battery predictive of *officer* success across the different commissioning sources, then constructing the AFOQT to include as many of the most critical non-GMA attributes as practically feasible should yield the greatest dividends. Making this happen, however, will require reconsidering the value-added from administering the flying specialty-specific tests to all prospective candidates, rather than administering them as part of an aviation-focused battery (e.g., Test of Basic Aviation Skills [TBAS]).

While there have been suggestions that one or more of existing non-AA subtests could be employed for making post-commissioning classification decisions (for non-rated officers), the ROI from this approach is expected to be comparatively less than the increments gained from adding new tests assessing higher-order abilities (i.e., fluid cognitive abilities-aptitudes), cross-functional skills (e.g., social-interpersonal skills), and/or personal qualities (e.g., personality). For example, consistent with the importance of predictor-criterion matching, we know that what increments in classification efficiency can be found over GMA from using similarly cognitively oriented tests (i.e., tests of crystallized cognitive abilities) mainly derives from tests targeting job family- or job-specific knowledge and skills (cf. Hunter, 1983; Olea & Ree, 1994). Of the five tests in question, only three measure job-relevant knowledge and skills, two of which (AI and IC) are content specific to flying specialties (the other is GS). Given this, there would seem to be comparatively less ROI from continuing to administer these tests to all prospective candidates than from incorporating alternatives (e.g., Figure Analogies, Missing Figures) assessing one or more of the predictor attributes cited in Table 7.[18]

Second, and related to preceding point, the fact that test-taking time is at a premium raises legitimate questions about practical dividends for the Air Force from continuing to maintain subtests measuring GMA when comparable alternative tests are available and their use would free-up time (approximately 60 minutes) to assess alternative predictor attributes not currently measured. While there have been concerns (e.g., differences in difficulty levels) raised about the most identifiable candidates – the SAT and the ASVAB (in particular the AFQT) – our review does not indicate that substituting one of these alternatives would significantly (and negatively) impact officer selection, as far as assessing GMA is concerned.[19] Differences in normative populations notwithstanding, earlier versions of the SAT and AFOQT-AA have been shown to be highly correlated (uncorrected *r*s = .80 to .85 for SAT Total scores) (Cowan et al., 1989; Diehl, 1986; Ree & Carretta, 1998; Ree, Carretta, & Earles, 2003).[20] When scores are corrected for relevant statistical artifacts (e.g., range restriction), the SAT and AFOQT-AA demonstrate a high degree of measurement- and construct-equivalence (i.e., corrected *r*s = .91-.95), as well as outcome equivalence, as the use of either leads to similar decisions on who is and is not selected (Ree & Carretta, 1998; Ree et al., 2003).[21] Consistent with this, in a large-scale sample of AFROTC cadets (Cowan et al., 1989), the SAT and AFOQT exhibited (a) comparable patterns of (uncorrected) criterion-related validities for different commissioning and post-commissioning performance criteria, and (b) similar patterns of correlations with other predictors (e.g., GPA) (see Table 10 and Figures 4 and 5).[22] Comparable results were observed for technical training performance – SAT scores correlated .39 with final course grades, while AFOQT-AA scores correlated .37 (Cowan et al., 1989, p. 8). In sum, the available information indicates that the AFOQT subtests targeting GMA and SAT (at least earlier versions) are conceptually and practically equivalent. Further, the newest version of the SAT could prove advantageous in its own right (cf. Breland, Kubota, Nickerson, Trapani, & Walker, 2004; Kobrin & Schmidt, 2005) since it now includes a writing sample test assessing written communication and critical thinking skills – all of which represent attributes expected to predict officer (commissioning and post-commissioning) success, which are not directly assessed by the current AFOQT.[23]

---

[18] One of the existing subtests, Table Reading (TR), could have potential for selecting non-flying specialty officers, since it potentially provides an assessment of Information Processing Aptitude (or Skills), although it is not likely to be predictive of global commissioning (training) criteria.

Similar findings have been made when comparing the ASVAB, specifically the Armed Forces Qualifications Test (AFQT), to the AFOQT subtests measuring GMA. Differences in factor structure notwithstanding, analyses have found that the ASVAB-AFQT correlates highly with earlier forms of the AFOQT-AA (corrected $r = .93$) (Carretta & Ree, 1995), as do subtests across the two batteries measuring the same attributes (Sperl et al., 1992). Further, earlier concerns about differences in test difficulty are no longer applicable with the ASVAB (specifically the AFQT), as DoD currently employs an operational computer-adaptive test (CAT) version of the ASVAB, which permits equal discrimination among test-takers across the ability range.[24] Finally, while we know of no direct comparison to AFOQT-AA, the AFQT has been found to be similarly predictive of technical training and job performance criteria for enlisted personnel and noncommissioned officers (NCOs) across the Services (cf. Campbell & Knapp, 2001; Welsh, Kucinkas, & Curran, 1990). Generally, between the SAT and ASVAB, the SAT is arguably preferable, as it (a) is comparatively more cost-effective (i.e., does not require Air Force or DoD administration), (b) is strongly predictive of success across academic and/or educational settings, (c) is currently used operationally in place of AFOQT-AA by at least one of the commissioning sources (AFROTC), and (d) expands the assessment of relevant predictor attributes at no cost to the Air Force, thereby freeing-up space for AFOQT to assess other critical attributes not measured by either test. Having said that, we return to this issue in greater depth in the next section, comparing alternative predictor attributes and assessment methods.[25]

---

[19] Additional study is warranted, as both the SAT and AFOQT have recently undergone significant revisions in content.

[20] (Uncorrected) correlations between equivalent composites scores were .76-.77 for Verbal and .71-.75 for Quantitative (Diehl, 1986; Ree et al., 1998).

[21] Provided one equates for differences in the distributional shape of the scores obtained from the two tests. Following this equating, the correlations between the SAT (Total) and AFOQT-AA is .97, which is comparable to the correlations observed when equating the new and former AFOQT composites (Gould et al., 2003).

[22] Because the AFOQT is normed to a population higher in ability, on average, than the SAT, it is not unreasonable to expect to see greater variability (and more potential for correlating with performance criteria) in AFOQT-AA scores than in SAT scores. One can see evidence for this by comparing their variability after adjusting for differences in scale – specifically, by computing their respective coefficient of variation (CV). In the Cowan et al. (1989) study, the CV for SAT (Total) was .1593276. In contrast, the CV for AFOQT-AA was nearly three times this number at .4952651. Therefore, even without adjusting for differences in variability, the pattern of correlations among AFOQT-AA score and SAT (Total) are very comparable.

[23] Experimental tests (e.g., Figure Analogies, Missing Figures) that potentially measure some of these attributes have been developed and initially piloted to, but are currently not in operational use.

Third, and as Table 10 demonstrates, while initial research on the SDI shows promise (e.g., Christal et al., 1997), there are reasons to be concerned about its operational use as a tool to select and potentially classify officers. Standard psychometric issues notwithstanding (e.g., retesting, stability of test scores, and so on), the most pressing has to do with the susceptibility of self-report measures, like the SDI, to possible compromise – resulting from intentional response distortion (i.e., faking) and coaching – and its implications for officer selection and classification. While research on the practical effects of faking and coaching (i.e., their effects on criterion-related validity) and its prevalence in operational selection contexts have been mixed (e.g., Barrick & Mount, 1996; Ellingson, Sackett, & Hough, 1999; Hough, 1998; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Mueller-Hanson, Heggestad, & Thornton, 2003; Ones & Viswesvaran, 1998; Rosse, Stecher, Levin, & Miller, 1998; White, Young, & Ramsey, 2001), there is a general consensus that (a) individuals can inflate or otherwise distort their responses to self-reports, (b) that with coaching or knowledge of the organization and/or job to which they are applying, they can do so in the desired direction, and (c) that faking and coaching can adversely impact the psychometric properties and scores of self-reports (Alliger & Dwight, 2000; Ellingson et al., 1999; Hough & Ones, 2001; McFarland & Ryan, 2000; Stark, Chernyshenko, Chan, Lee, & Drasgow, 2001; Viswesvaran & Ones, 1999; Vasilopoulos, Reilly, & Leaman, 2000; White et al., 2001).[26] In our estimation, this issue warrants special attention because of the practical effects it has had on prior military-directed efforts to develop and operationally implement self-report personality inventories for selection and classification purposes. During the 1980s and 1990s, for example, the Army expended a considerable amount of resources and time developing and studying implementation issues for a comparable self-report personality measure, the Assessment of Background and Life Experiences (ABLE), and its use in pre-enlistment screening and post-enlistment selection and classification decisions. While the results of several

---

[24] Currently, the CAT-ASVAB is administered about 400,000 times annually. Additionally, there are several research efforts presently underway to pilot and evaluate an Internet-based version (iCAT-ASVAB).

[25] We return to this issue in greater depth in the next section comparing alternative predictors and assessment methods.

[26] Ultimately, the practical implications (and seriousness) of faking and coaching effects could depend heavily on the SDI+ is used. For example, potential faking and coaching effects are more likely to exert a serious (negative) impact if the SDI+ is used to rank-order officer candidates than if employed as an initial screen to ensure candidates meet some minimum requirement (i.e., cut score).

Figure 5. Comparison of SAT (Total) and AFOQT-AA by Performance Criteria

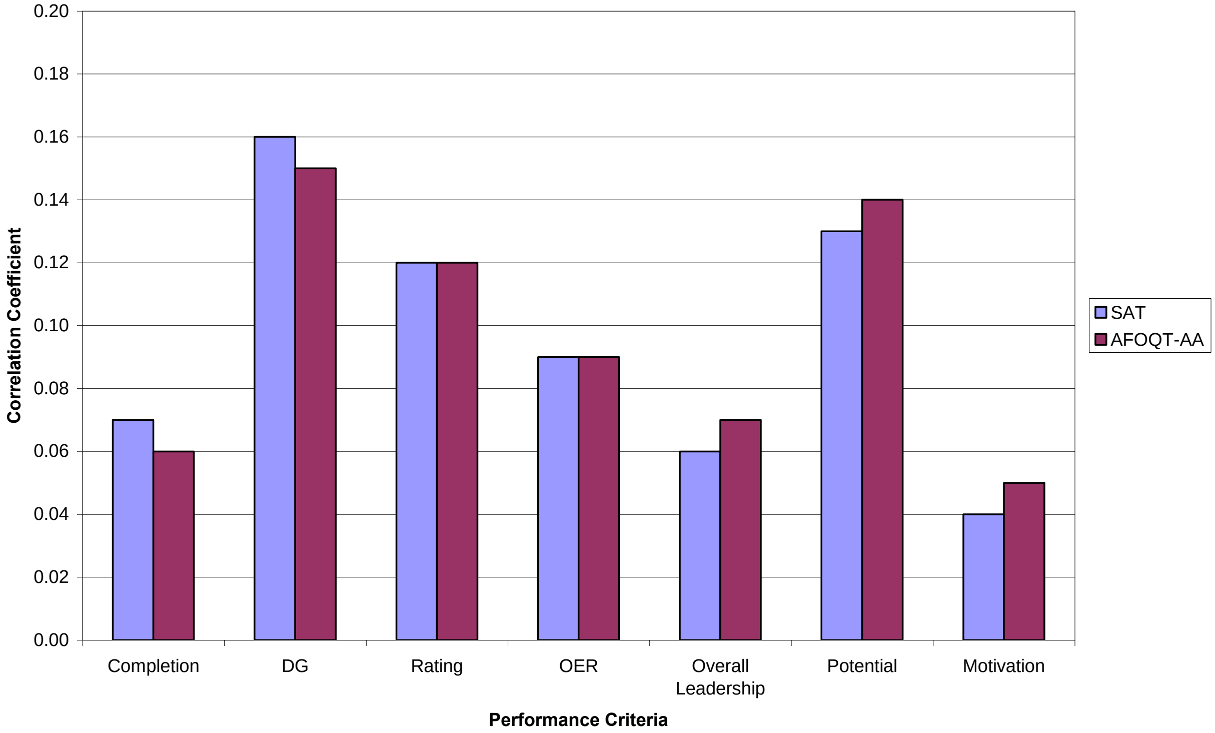
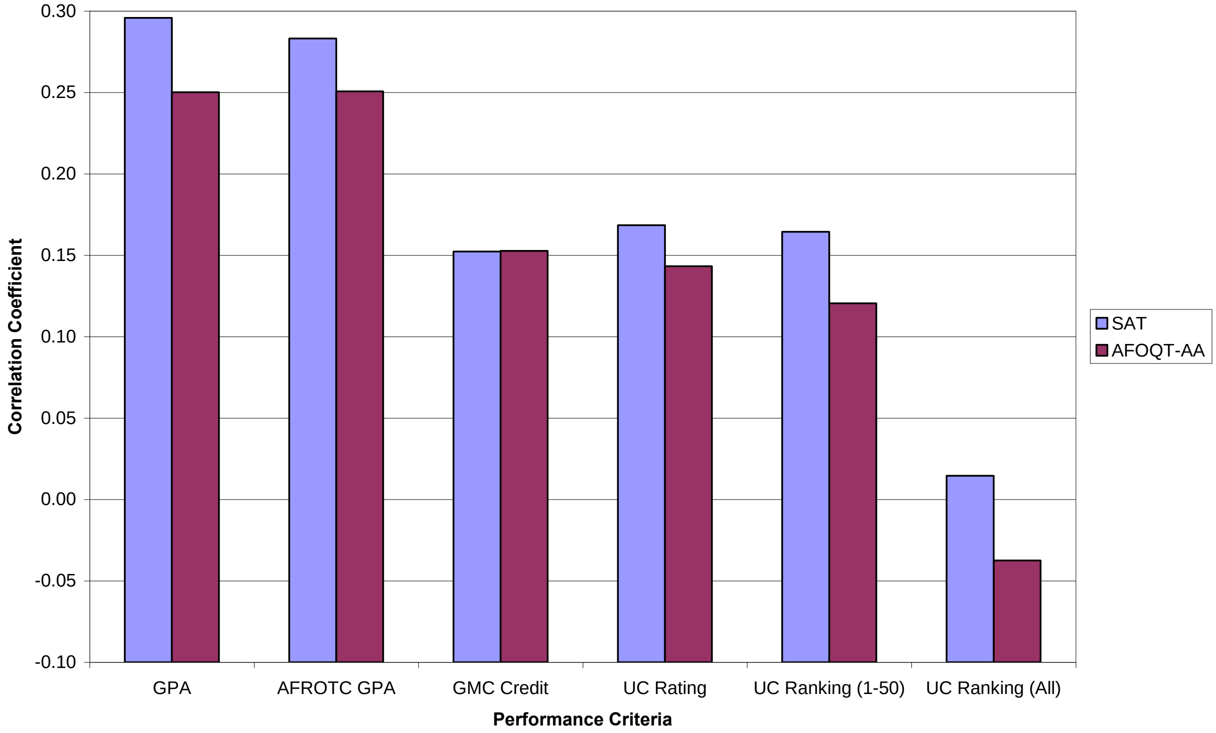


Figure 6. Comparison of SAT (Total) and AFOQT-AA by AFROTC Predictor



of these studies were promising, the ABLE was ultimately abandoned, at least as a pre-screen or selection tool. This came about because even if effective controls for faking and coaching could be implemented, legitimate concerns remained that the ABLE’s highly transparent and nonverifiable nature would make it a potential and continued target of criticism (White et al., 2001). As a result, the Army now employs alternative assessment methods, such as multidimensional forced-choice item formats (e.g., the Assessment of Individual Motivation [AIM], Tailored Adaptive Personality Assessment System [TAPAS]), to assess personality and similar non-cognitive attributes. Given the high stakes and visibility associated with officer selection, the Air Force can be expected to face comparable pressures and concerns from stakeholders, which could preclude the operational use of the SDI+ as a selection and classification tool.

Step 3: Recommendations for Closing Critical Gaps

The following recommendations suggest actions the Air Force can take to close the gaps identified from our evaluation. As the sister Services and other organizations are experiencing similar challenges, several of these recommendations reflect actions they are taking to overcome these gaps. Recommendations are ordered in terms of priority (short-term versus long-term, i.e., prior to iteration of the next AFOQT) and by selection system component (strategy, content, and method).

Short-Term

Selection System Strategy

- *Recommendation 1. Develop and implement a competency (performance) model – to include retention – and align officer selection systems (and training) accordingly, so that performance requirements measurably drive selection.* As summarized earlier, the Air Force has already developed and taken steps to implement an officer competency (performance) model, but the model needs to be expanded to ensure that it accomplishes its desired objectives. Specifically, expanding the current model will require specifying (a) behavioral exemplars for each competency, primarily for training performance management purposes, and (b) linkages between competencies and relevant predictor attributes (e.g., abilities-aptitudes, cross-functional skills, and personal qualities). To implement this model operationally, ―value chain” analysis could prove useful, as it represents an effective means for articulating the linkages among Force needs and goals, such as those flowing from the *Personnel Strategic Plan*, to specific competencies (performance), training needs, and work environment needs, so that everyone is on the ―same page.” From there, the current selection systems (and their components), related processes (e.g., training), and so on can be systematically reviewed and realigned against relevant performance requirements.

- *Recommendation 2. Implement a sufficiently comprehensive criterion assessment system that enables the measurement of desired results, so that the success of officer selection systems (and their components) can be evaluated and improved (as needed).* Ideally, this system will include metrics that reflect (a) officer competency (performance) requirements (commissioning and post-commissioning), (b) retention-related imperatives, and (c) other critical outcomes, where feasible, at multiple individual levels-of-analysis (e.g., team or unit effectiveness). Equally as important, it would include diversity-based metrics derived from clear and sufficient specifications of officer diversity requirements. Having such a system will enable the Air Force to (a) periodically track the success of officer selection systems (and their components) in meeting Air Force needs, (b) provide stakeholders the data needed to systematically evaluative the ROI of different predictors, assessment methods, design choices, and so on, and (c) to prescriptively model how best to improve officer selection (e.g., the relative emphasis to place on officership versus technical expertise). The Army is currently working on such a system (minus diversity metrics), albeit geared towards enlisted personnel and noncommissioned officers (NCOs) within the context of promotion, as part of its Army21 transformation efforts.

- *Recommendation 3. Investigate and take advantage of ways in which current (or proposed) assessment tools used for officer selection could be integrated with recruitment and training (and vice versa) to complement and maximize the Air Force's ROI in these processes*. Ultimately, integrating officer assessment tools (as a set) programmatically across different HR functions (e.g., recruitment, selection, training, etc.) could pay greater ROI to the Air Force and the other services than continuing to use or consider these assessments individually and exclusive to a specific function. For example, more and more organizations are using shortened versions of operational selection tests as a means to prescreen and facilitate the –selecting out” of less desirable or qualified candidates. Doing so is advantageous because it cuts down on subsequent selection costs without having to significantly raise minimum eligibility requirements in a way that inadvertently promotes a negative image of the organization. Similarly, organizations are exploring methods by which data and scores obtained during the selection process can be leveraged to supply targeted, diagnostic feedback to prospective candidates (and the organization), so that each can maximize future opportunities and staffing decisions (e.g., training). Perhaps equally as important, many of the assessment tools currently in use (or under consideration) could be modified to effectively support outreach efforts to promote hard-to-fill specialties (e.g., pilots-navigators) or military service generally to younger high school students to increase the future applicant pool. As an example, DMDC is currently funding a couple of initiatives, the ASVAB web site and Career Exploration Program (that includes a career interest inventory), in this regard. Finally, these platforms could provide an effective, reasonably cost-effective means for experimenting with new assessment (or selection) technologies.

Selection System Content

- *Recommendation 4. Partner with AFRS and AFROTC (and USAFA) to standardize the specification and conceptualization of predictor attributes used for officer selection*. For example, standardize what is meant by –officer potential.” Standardizing *what* predictor attributes (i.e., their specification and what they mean) are assessed and selected on is desirable, even if *how* these attributes are measured and weighted operationally differs by commissioning source.[27] Doing so is advantageous because it would promote consistency across and within the different commissioning sources. At a minimum, such consistency would ensure that the predictor attributes actually assessed (and selected) on are the same and in line with Air Force competency (performance) requirements, thereby minimizing possible errors that could detract from their ROI and enhancing each source's ability to meet relevant Force needs. In addition, standardization, coupled with a more structured assessment, could be accomplished for a reasonable amount of effort and expense.

- *Recommendation 5. Maximize coverage of predictor attributes not currently assessed whose addition could (and would) significantly increment the prediction of officer success and promote Force diversity*. Ideally, officer selection will seek to maximize coverage of those predictor attributes *most* critical to officer success (commissioning and post-commissioning). While we would not suggest that the Air Force should (and can) measure everything, there are a number of predictor attributes not currently covered that could prove advantageous to officer selection. As many organizations within and outside the military are finding there is a strong need to assess a greater number and range of predictor attributes than in the past (e.g., Borman et al., 2003; Ford et al., 2000; Ilgen & Pulakos, 1999; Pearlman, 2003; Pearlman & Barney, 2000; Sager et al., 2004). Beyond optimizing prediction, expanding the predictor space would prove useful for promoting a diverse Force – both in terms of demography and capabilities (i.e., knowledge, skills, and abilities, KSAs). Additionally, many of these attributes – Cultural Tolerance, Teamwork Skills, and so on – conform with and can be reasonably expected to facilitate the Air Force's efforts to meet emerging DoD-wide initiatives and Joint Force needs (e.g., officers with foreign language and cultural proficiency). While data relating a number of these predictor attributes to relevant officer competency (performance) requirements (and each other) is limited, we make a systematic attempt in the next section to narrow this list down to those most likely to prove promising.

Selection System Method, Design, and Implementation

- *Recommendation 6. Meaningfully and measurably separate selection for officership from selection for technical expertise*. By employing a (additive) compensatory model that considers officer potential and technical potential simultaneously, the current officer selection systems create potential ―tradeoffs” between selecting ―good” officers and ―good” technical specialists, specifically pilots-navigators, which could be avoided by using an alternative selection model. One potential solution, albeit not the only one, would be a multiple hurdle-type model – whereby prospective candidates are screened for a sufficient level of officer potential first then selected based on predictors relevant to technical performance (or technical training success). Whatever model is employed, the point is that explicitly and measurably separating officer selection from specialist selection would be advantageous, as it enables selection decision-makers to make an informed decision as to how best to weight potential tradeoffs (as needed). While there may be practical concerns about the use of particular models (e.g., multiple hurdle), and their impact on meeting end strength requirements, we suspect that future Force Shaping and comparable initiatives will facilitate the employment of these models. Generally speaking, the best long-term solution may ultimately be to adopt a flexible (rather than fixed) approach, whereby the selection model changes as the Air Force's needs change.

---

[27] Ultimately, and as discussed previously, how much standardization in terms of weighting, etc. is needed will depend on several factors, including the degree to which the different commissioning sources respectively serve differing Air Force's needs. Implementing the preceding recommendations (under Strategy) should clarify these factors.

- *Recommendation 7. Partner with AFRS and AFROTC to standardize and structure existing measures used in officer selection*. As discussed previously, several of the assessment measures most strongly contributing to AFROTC and OTS officer selection decisions (e.g., RSS, recruiter/interviewer evaluations) produce criterion-related validities that, at least empirically, do not support these weightings. All other things being equal, the most immediate explanation, other than the absence of substantive and clear specifications of the predictor attributes they are meant to measure, points to the fact that they generally reflect unstandardized (or unstructured) assessments. Adding (some) standardization and structure to these assessments (e.g., behaviorally-anchored ratings scales (BARS) congruent with emerging competency requirements) may not necessarily require extensive resources or the development of completely ―new" assessments, thereby paying reasonable dividends to the Air Force.

- *Recommendation 8. Partner with AFRS to systematically review and evaluate the continued use of specific assessments (e.g., letters of recommendation) for OTS officer selections*. OTS employs several assessments whose ROI appears limited, as they are not meaningfully correlated with commissioning or post-commissioning performance criteria and/or are not heavily used by selection boards when making decisions (Cowan et al., 1990). In several instances, alternative measures (e.g., SDI+) are potentially available that could provide more valid and efficient assessments of critical predictor attributes than the measures currently in use, such that the tradeoffs (in terms of costs) would not be substantial and would likely favor the more standardized, valid assessment tool.

- *Recommendation 9. Investigate implementation issues and related concerns (e.g., faking) that might preclude the use of the SDI+ as an officer selection and classification tool*. As mentioned previously, while the SDI+ shows promise as an officer selection tool, historically self-report measures have suffered from implementation concerns and issues (e.g., faking and coaching) uncommon to traditional aptitude-ability tests. While not to trivialize the importance or need for systematic study of more traditional issues (e.g., test-retest reliability), effectively addressing implementation issues has and continues to remain the main stumbling block to the operational use of personality in military selection and classification (e.g., White et al., 2001).

Long-Term

Selection System Strategy

- *Recommendation 10. Ensure there is a formal, programmatic system in place for sustaining and maintaining linkages between officer selection systems (and their components) and potential changes to competency (performance) requirements and other critical imperatives (e.g., retention).* Doing so would ensure and sustain (over time) the right alignment and consistency between AFROTC and OTS in their selection systems. A significant part of this programmatic system would include the components previously mentioned (e.g., a comprehensive officer competency-performance model, meaningful performance assessment, etc.). In addition, it may be advantageous to formalize this process in a mechanism comparable to, or meaningfully incorporated into, the Commissioning Education Memorandum of Understanding (CEMU), which serves to focus and coordinate officer curriculum and training across commissioning sources.

Selection System Content

- *None.*

Selection System Method, Design, and Implementation

- *Recommendation 11. Systematically study and investigate use of the new SAT Reasoning Test in place of the GMA-specific subtests in current AFOQT and/or feasibility of moving flying specialty-specific tests to a specialty-specific battery.* Provided that the primary goal of the AFOQT is to predict *officer* success (commissioning and post-commissioning), expanding it to incorporate other critical predictor attributes not currently covered will likely require reductions and changes in the current battery. At present, most of this battery consists of (a) subtests assessing GMA (and constituent specific aptitudes-abilities reflecting Verbal and Quantitative Ability), or (b) subtests targeting abilities-aptitudes or basic skills exclusive to flying specialties (pilot-navigator). While we expect that GMA will continue to be relevant to the selection of officers and the specific aptitudes-basic skills for pilots-navigators, there are arguably viable, alternative means for assessing these attributes respectively that would free-up test-taking time on the AFOQT. For example, from our review of available research, replacing the AFOQT GMA-focused subtest scores with SAT Reasoning Test scores is not expected to practically impact officer quality. Nevertheless, both tests recently underwent some substantial revisions, which could potentially impact estimates of their comparative ROI – although more so probably for the SAT than the AFOQT. Most importantly, we found no research directly speaking to its implications for the OTS population, where SAT scores will be four or more years old at the time a candidate applies for commissioning (i.e., there is the possibility of maturation effects). While there is research indicating that SAT scores can predict career-related outcomes more than 10 years later (e.g., Wai, Lubinski, & Benbow, 2005), and these scores combined with cumulative college GPA may be sufficient to compensate for any loss in predictive efficacy from not obtaining a more recent assessment of GMA, at present we do not know that definitively. As for the comparative pros and cons of the SAT relative to other readily accessible standardized tests (e.g., ASVAB), we return to this issue in greater depth in the next section.

- *Recommendation 12. Investigate possibilities for a joint-service (junior) officer selection battery*. Generally speaking, across the sister Services, the junior officer competency (performance) requirements are the same (see Table 5). Because of this, maintaining ―homegrown,” Service-specific measures, specifically of more distal predictor attributes (e.g., personality), is probably not needed since the attributes assessed will be the same and are generally not context-sensitive – although, where there could be substantive differences is in how the different predictor attributes are weighted to form composites scores. Additionally, with continued budgetary pressures, maintaining these tests within each service long-term is arguably not cost-effective. As is currently done with the ASVAB, practical issues and potential concerns – such as between-service population differences in officer characteristics – can be successfully addressed by generating Service-relevant norms and using testing technologies (e.g., CAT) that take into account differences in test (item) difficulty. While readily accessible standardized tests, such as the SAT Reasoning Test, provide assessments of GMA (and other cognitively-oriented attributes) comparable to and potentially richer than the AFOQT, the tests are ultimately non-DoD owned and therefore may not always reflect DoD needs and priorities. Further, and most importantly, they do not cover a number of the cross-functional skills and/or personal qualities required to be an officer in the U.S. military. For this reason, we recommend that the Air Force continue to follow developments in the ASVAB (e.g., iCAT-ASVAB), which if nothing else, could potentially provide an infrastructure and platform for administering a future joint-service officer selection battery.

Towards an Officer Assessment “Toolkit” for the 21st Century: Identification and Evaluation of Alternative Predictors and Assessment Methods

As other organizations are discovering, and as suggested in the preceding section, the Air Force could potentially reap substantial benefits from modifying how officers are currently selected and assessed. To facilitate and inform possible changes, we considered and evaluated several alternative predictors and assessment methods with an eye towards formulating recommendations on an officer selection “toolkit” for the 21st Century. Formulating this “toolkit” consisted of three steps: (a) identifying what attributes to target for selection not currently covered (or covered well) by the current selection systems, (b) comparing and evaluating different assessment methods for assessing targeted attributes, and (c) generating specific, actionable recommendations for constructing an officer selection “toolkit” to serve the 21st Century Air Force. Figure 6 summarizes this process.

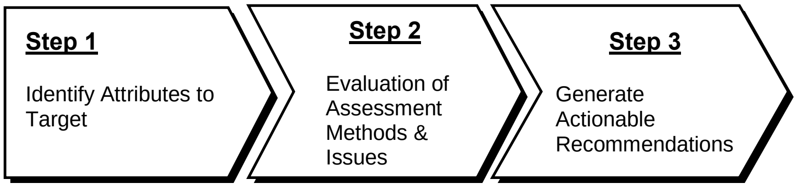

Step 1
Identify Attributes to Target
Step 2
Evaluation of Assessment Methods & Issues
Step 3
Generate Actionable Recommendations

Figure 7. Steps in Formulating an Officer Selection “Toolkit”

Step 1: Identifying Predictor Attributes to Target for Selection

What We Did

As presented earlier (see Table 7), the number and categories of predictor attributes that have been conceptually and/or empirically related to leader performance is fairly extensive. Practically, this presents the Air Force with a challenging array of attributes to (potentially) measure and include in its officer selection systems. To meaningfully narrow down this list, we conducted a comprehensive literature review. Relevant sources reviewed included meta-analyses (e.g., Judge et al., 2002; Judge et al., 2004), narrative literature reviews (e.g., Avolio, Sosik, Jung, & Berson, 2003; Wong et al., 2003; Yukl, 2005; Zaccaro, Kemp, & Bader, 2004), published research studies (e.g., Bass, Avolio, Jung, & Berson, 2003; Chan & Drasgow, 2001; Connelly, Gilbert, Zaccaro, Threlfall, Marks, & Mumford, 2000; Pulakos et al., 2002), technical reports, conference papers, and workshop proceedings (e.g., Ford et al., 2000; Horey, Morath, McGonigle, Cronin, & Wilson, 2005; Horey et al., 2004; Knapp et al., 2004; Pearlman, 2003), produced in the military and civilian sectors on predictors of organizational leadership and/or facets of organizationally-relevant criteria (e.g., retention) reflecting one or more of the Air Force’s competency (performance) requirements specified earlier. Additionally, where feasible, we supplemented this information with expert judgments. To better understand potential attribute-performance linkages so as to ensure that the final list contained those attributes *most* critical to target (and as a means to guide subsequent research), we carefully matched predictors to their relevant competency (performance) domain. In making our final selections, we focused on the following criteria:

- Expected and/or demonstrated ability to meaningfully predict one or more of the Air Force's competency (performance) requirements;
- Potential to add incrementally to prediction beyond general mental ability (GMA) (i.e., AFOQT). Generally speaking, these would be (a) higher-level cognitive aptitude-abilities (i.e., fluid cognitive abilities-aptitudes, such as creative-divergent thinking, critical thinking, problem-solving), (b) cross-functional skills (e.g., self-management skills, social-interpersonal skills), and (c) personal qualities reflecting social-motivational attributes (e.g., adaptability, integrity);
- Degree to which predictor attributes were trainable and/or were desired at entry. For example, personal qualities (e.g., personality) are generally regarded as not trainable, whereas cross-functional skills are. However, having some level of proficiency on select cross-functional skills are likely to be desired at entry, so as to reduce and minimize the time and resources needed for training. As evident by Force Shaping and other Force Management efforts, issues of trainability and desired at entry are expected to increase in importance as junior officers' level of responsibility and rate of promotability concomitantly increases to meet Air Force needs.

Findings

Table 13 shows those predictor attributes that emerged as being most consistently and strongly related to the Air Force's performance (competency) requirements. To compliment this table, Table 14 organizes recent meta-analytic research (e.g., Judge et al., 2002; Judge et al., 2004) relating select attributes to various leader emergence (i.e., who becomes a leader) and effectiveness (i.e., how well the leader or team/group s/he leads performs) criteria. As can be seen from Table 14, some of these attributes (i.e., GMA and the Big Five personality traits) have received greater empirical coverage than others, specifically as it applies to the prediction of organizational leadership. Nevertheless, as suggested by the table, and demonstrated by recent research (e.g., Chan & Drasgow, 2001; Connelly et al., 2000; Knapp et al., 2004; Pulakos et al., 2002), there is still much to potentially gain in terms of incremental validity from considering these other predictors. Especially if one of the goals of selection is to maximize the prediction of future, on-the-job performance, or a comparable effectiveness criterion (e.g., team or unit performance), which tends to be more complex and multidimensional than technical training performance and where specific attributes have more of an opportunity to exhibit their potential incremental validity (e.g., Olea & Ree, 1994; Ree et al., 1994).

While we are not suggesting that the Air Force can (and should) attempt to measure all of the predictor attributes listed, those identified from our review represent the best choices. Ideally, and consistent with the Air Force's emphasis on the —whole person" concept, an assessment battery would measure as many of these performance-relevant attributes as possible. Tables 15 and 16 provide a summary of the degree to which the current selection system – or components under serious consideration, such as the SDI+ assesses these attributes. Generally, there are comparatively few measures that directly and sufficiently tap one or more of these attributes. However, several measures could provide relevant information on these attributes, particularly if considered in the aggregate; although, the degree to which this is actually the case, is an open empirical question.

Three additional substantive points should be taken from this review. First, and as implied in Table 13, we expect that the attributes associated with particular performance (competency) requirements will for various substantive and methodological reasons, cluster and covary together (Hogan & Kaiser, 2005; Kirkpatrick & Locke, 1991; Yukl, 2005; Zaccaro et al., 2004). This raises questions about the ROI of measuring most (or all) attributes linked to a given performance (competency) requirement. For example, we generally know that one's development of a skill – or skills set (e.g., social-interpersonal skills) – is, to a measurable extent, predicted by his/her possession of related distal attributes (e.g., ability-aptitude, personality) (Borman et al., 2003; Campbell et al., 1993; Connelly et al., 2000; Pearlman & Barney, 2000; Zaccaro et al., 2004). Therefore, while conceptually there may be a meaningful difference between affiliated distal attributes and skills, from a predictive standpoint measuring both may not be needed and could prove costly. Having said that, however, using composites (or profiles) of multiple attributes can, in and of themselves, produce greater predictive efficacy than considering any of the attributes separately (e.g., Davison & Davenport, 2002; Judge et al., 2002). For this reason, we recommend, as much as feasible, assessing multiple conceptually related attributes. Practically speaking, this is important since different assessment approaches are more or less effective in measuring different attributes.

Second, we see great potential for focusing on cross-functional skills such as social-interpersonal and teamwork skills. Others have made a similar recommendation (e.g., Drasgow, 2003; Pearlman, 2003; Pearlman & Barney, 2000). Assessing cross-functional skills is advantageous for two reasons. First, from a causal perspective, knowledge and skills (or skill sets) are more proximal to performance and thereby tend to be more predictive than distal attributes (e.g., Borman, White, Pulakos, & Oppler, 1991; Borman, Hanson, Oppler, Pulakos, & White, 1993; Ree, Caretta, & Teachout, 1995). Practically, the kinds of intra- and inter-personal skills identified by our review represent those that the population targeted by officer assessment (ages 18 on) can be reasonably expected to have developed sufficient proficiency in for purposes of making a valid predictive assessment (Campbell, 2003; Hogan & Kaiser, 2005). The second advantage to an emphasis on skills is that they offer a potentially more efficient and valid way of assessing more socially-motivationally valenced attributes than traditional methods, such as self-reports (McDaniel, Finnegan, Morgeson, Campion, & Braverman, 2001; Knapp et al., 2004; Pearlman, 2003), although research targeting test compromise issues is limited (Hooper, Cullen, & Sackett, 2006). While skills assessment carries its own challenges, in the long-term the benefits are likely to outweigh the costs.

Third, one potential and significant barrier to the value of several of these attributes as predictors is that currently we do not have substantive specifications of exactly what they are, which makes valid assessment difficult. This is especially true for those attributes that are multi-faceted or multi-dimensional (e.g., social-interpersonal skills; social-emotional intelligence, etc.). ―Adaptability" is a highly visible example. While there have been systematic efforts to specify

Table 13. Predictor Attributes Needed for 21st Century Officer Success

Performance Imperative	Representative Competencies	Relevant Predictor Attributes
Personal Leadership	Exercise Sound Judgment	• Analytical Reasoning (Skills) • Creative-Divergent Thinking (Skills) • Critical Thinking (Skills) • Information Processing & Analysis (Skills) • Problem-Solving (Skills)
	Adapt and Perform Under Pressures	• Adaptability • Core Self-Evaluation (Emotional Stability) • Energy Level-Potency • Self-Management Skills
	Lead Courageously	• Achievement Orientation-Initiative • Conscientiousness-Integrity • Self-Management Skills
	Assess Self	• General Mental Ability (GMA) • Cognitive-Integrative Complexity • Core Self-Evaluation (Emotional Stability) • Motivation to Learn-Development Orientation • Self-Directed Learning & Development Skills
	Foster Effective Communication	• Information Processing & Analysis (Skills) • Social-Interpersonal-Teamwork Skills • Verbal-Written Communication Skills
Leading People/ Teams	Inspire Trust; Drive Performance; Influence Through Win/Win Solutions; Mentor and Coach; Promote Collaboration & Teamwork; Partner to Maximize Results	• Adaptability • Conscientiousness-Integrity • Motivation to Lead-Socialized Power Motive • Openness to Experience-Cultural Tolerance • Service-Team Orientation • Social-Interpersonal-Teamwork Skills
Leading the Institution	Embrace Change and Transformation; Drive Execution	• Adaptability • Achievement Orientation-Initiative • Creative-Divergent Thinking (Skills) • Motivation to Lead-Socialized Power Motive
Retention	Person-Role Fit; Person-Organization (P-O) Fit	• Motivation to Lead-Socialized Power Motive • Service-Team Orientation

Table 14. Summary of Meta-Analytic Validity Estimates Relating Select Attributes to Leader Emergence and Effectiveness

			Effectiveness (Performance)		
Attribute	**Overall**	**Emergence**	Global	Transformational	Transactional
General Mental Ability (GMA)	$\bar{r}$ = .17, ρ_{pc} = .21, ρ_t = .27 *N* = 40,652; *k* = 151 Judge et al., 2004	ρ_{pc} = .19, ρ_t = .25 *k* = 65 Judge et al., 2004	ρ_{pc} = .15, ρ_t = .18 *k* = 34 Judge et al., 2004; individual-level		
			ρ_{pc} = .19, ρ_t = .22 *k* = 26 Judge et al., 2004; group-level		
Agreeableness	$\bar{r}$ = .06, ρ_{pc} = .08 *N* = 9,801; *k* = 42 Judge et al., 2002; all settings	ρ_{pc} = .05 *k* = 23 Judge et al., 2002; all settings	ρ_{pc} = .21 *k* = 19 Judge et al., 2002; all settings	$\bar{r}$ = .10, ρ_{pc} = .14 *N* = 3,916; *k* = 20 Bono & Judge, 2004	$\bar{r}$ = .03, ρ_{pc} = .04 *N* = 1,564; *k* = 7 Bono & Judge, 2004
	ρ_{pc} = -.04 *k* = 11 Judge et al., 2002; government/military settings				
Conscientiousness	$\bar{r}$ = .20, ρ_{pc} = .28 *N* = 7,510; *k* = 35 Judge et al., 2002; all settings	ρ_{pc} = .33 *k* = 17 Judge et al., 2002; all settings	ρ_{pc} = .16 *k* = 18 Judge et al., 2002; all settings	$\bar{r}$ = .10, ρ_{pc} = .13 *N* = 3,516; *k* = 18 Bono & Judge, 2004	$\bar{r}$ = -.09, ρ_{pc} = -.11 *N* = 1,564; *k* = 7 Bono & Judge, 2004
	ρ_{pc} = .17 *k* = 6 Judge et al., 2002; government/military settings				

Note. $\bar{r}$ = sample-size weighted mean observed validities; ρ_{pc} = validities corrected for unreliability in both predictor and the criterion; ρ_t = validities corrected for range restriction and unreliability in both predictor and criterion; *N* = combined (total) sample size; *k* = number of correlations.

Table 14. Summary of Meta-Analytic Validity Estimates Relating Select Attributes to Leader Emergence and Effectiveness (cont'd)

Attribute	**Overall**	**Emergence**	**Effectiveness (Performance)**		
			Global	Transformational	Transactional
Conscientiousness (cont'd)					
Achievement Motivation	$\bar{r}$ = .23, ρ_{pc} = .35 N = 4,625; k = 16 Judge et al., 2002; all settings				
Dependability	$\bar{r}$ = .18, ρ_{pc} = .30 N = 5,020; k = 16 Judge et al., 2002; all settings				
Emotional Stability	$\bar{r}$ = .17, ρ_{pc} = .24 N = 8,025; k = 48 Judge et al., 2002; all settings	ρ_{pc} = .24 k = 30 Judge et al., 2002; all settings	ρ_{pc} = .22 k = 18 Judge et al., 2002; all settings	$\bar{r}$ = .15, ρ_{pc} = .17 N = 3,380; k = 18 Bono & Judge, 2004	$\bar{r}$ = -.04, ρ_{pc} = -.05 N = 1,627; k = 8 Bono & Judge, 2004
	ρ_{pc} = .23 k = 12 Judge et al., 2002; government/military settings				
Self-Esteem	$\bar{r}$ = .14, ρ_{pc} = .19 N = 7,451; k = 9 Judge et al., 2002; all settings				

Note. $\bar{r}$ = sample-size weighted mean observed validities; ρ_{pc} = validities corrected for unreliability in both predictor and the criterion; ρ_t = validities corrected for range restriction and unreliability in both predictor and criterion; *N* = combined (total) sample size; *k* = number of correlations.

Table 14. Summary of Meta-Analytic Validity Estimates Relating Select Attributes to Leader Emergence and Effectiveness (cont'd)

Attribute	**Overall**	**Emergence**	**Effectiveness (Performance)**		
			Global	Transformational	Transactional
Extraversion	$\bar{r}$ = .22, ρ_{pc} = .31 N = 11,705; k = 60 Judge et al., 2002; all settings	ρ_{pc} = .33 k = 37 Judge et al., 2002; all settings	ρ_{pc} = .24 k = 23 Judge et al., 2002; all settings	$\bar{r}$ = .19, ρ_{pc} = .24 N = 3,692; k = 20 Bono & Judge, 2004	$\bar{r}$ = -.07, ρ_{pc} = -.09 N = 1,310; k = 6 Bono & Judge, 2004
	ρ_{pc} = .16 k = 10 Judge et al., 2002; government/military settings				
Dominance	$\bar{r}$ = .24, ρ_{pc} = .37 N = 7,692; k = 31 Judge et al., 2002; all settings				
Sociability	$\bar{r}$ = .24, ρ_{pc} = .37 N = 5,827; k = 19 Judge et al., 2002; all settings				
Openness to Experience	$\bar{r}$ = .16, ρ_{pc} = .24 N = 7,221; k = 37 Judge et al., 2002; all settings	ρ_{pc} = .24 k = 20 Judge et al., 2002; all settings	ρ_{pc} = .24 k = 17 Judge et al., 2002; all settings	$\bar{r}$ = .11, ρ_{pc} = .15 N = 3,887; k = 19 Bono & Judge, 2004	$\bar{r}$ = .03, ρ_{pc} = .04 N = 1,564; k = 7 Bono & Judge, 2004
	ρ_{pc} = .06 k = 6 Judge et al., 2002; government/military settings				

Note. $\bar{r}$ = sample-size weighted mean observed validities; ρ_{pc} = validities corrected for unreliability in both predictor and the criterion; ρ_t = validities corrected for range restriction and unreliability in both predictor and criterion; *N* = combined (total) sample size; *k* = number of correlations.

Table 14. Summary of Meta-Analytic Validity Estimates Relating Select Attributes to Leader Emergence and Effectiveness (cont'd)

Attribute	Overall	Emergence	Effectiveness (Performance)		
			Global	Transformational	Transactional
(Internal) Locus of Control	$\bar{r}$ = .08, ρ_{pc} = .13 *N* = 2,347; *k* = 15 Judge et al., 2002; all settings				
Self-Monitoring	$\bar{r}$ = .18, ρ_{p} = .21 *N* = 2,777; *k* = 23 Day et al., 2002				

Note. $\bar{r}$ = sample-size weighted mean observed validities; ρ_p = validities corrected for predictor unreliability; ρ_{pc} = validities corrected for unreliability in both predictor and the criterion; ρ_t = validities corrected for range restriction and unreliability in both predictor and criterion; *N* = combined (total) sample size; *k* = number of correlations.

Table 15. Coverage of Recommended Predictor Attributes by AFROTC Predictor Measures (Current and Proposed)

	AFROTC Predictor Measures (Current and Proposed)										
	PSP Order of Merit (OM)				*Self-Description Inventory (SDI+)*						
Recommended Predictor Attributes	RSS	GPA	AFOQT AA	PFT	Agree	Consc	Emot	Extra	Open	SO	TO
Cognitive-Integrative Complexity		O	O								
Creative-Divergent Thinking (Skills)											
Critical Thinking (Skills)		O									
General-Analytical Reasoning (Skills)		O	O								
Information Processing (Skills)			X								
Problem-Solving (Skills)		O	O								
General Mental Ability		O	X								
Verbal-Written Communication Skills		O	X								
Achievement Orientation	O	O				O					
Adaptability	O				O	O	O	O			O
Conscientiousness-Integrity	O	O			O	X	O				
Core Self-Evaluation (Emotional Stability)	O	O				O	X				
Energy Level-Potency	O					X		X			
Motivation to Lead	O				O			O		O	O
Motivation to Learn		O				O			O		
Openness to Experience-Cultural Tolerance					O				X		O
Service-Team Orientation	O									X	X
Self Directed Learning & Development Skills		O				O			O		
Self-Management Skills	O	O				O	O				
Social-Interpersonal Skills	O				O			O	O	O	O

Note. O = Expect to correlate with recommended predictor attribute(s). X = Directly (and sufficiently) measures one or more components of recommended predictor attribute(s).

Table 16. Coverage of Recommended Predictor Attributes by OTS Predictor Measures (Current)

	OTS Predictor Measures (Current)									
	Education-Aptitude		*Experience*					*Potential-Adaptability*		
Recommended Predictor Attributes	GPA	AFOQT-AA	Work Exp-NM	Work Exp-M	NM-Awards	NM-Achv	Extra Act	Eval-Overall	Eval-Com Skills	Rec Letter
Cognitive-Integrative Complexity	O	O								
Creative-Divergent Thinking (Skills)										
Critical Thinking (Skills)	O									
General-Analytical Reasoning (Skills)	O	O								
Information Processing (Skills)		X								
Problem-Solving (Skills)	O	O								
General Mental Ability	O	X			O	O				
Verbal-Written Communication Skills	O	X						O	X	
Achievement Orientation	O			O	O	O		O		
Adaptability			O	O	O	O	O	O		
Conscientiousness-Integrity	O			O	O	O	O	O		
Core Self-Evaluation (Emotional Stability)	O			O	O	O		O		
Energy Level-Potency			O	O	O	O	O			
Motivation to Lead								O		
Motivation to Learn	O							O		
Openness to Experience-Cultural Tolerance			O	O	O	O	O			
Service-Team Orientation					O	O	O			
Self Directed Learning & Development Skills	O				O	O	O			
Self-Management Skills	O			O	O	O	O			
Social-Interpersonal Skills			O	O			O	O		

Note. O = Expect to correlate with recommended predictor attribute(s). X = Directly (and sufficiently) measures one or more components of recommended predictor attribute(s).

what “adaptability” as a predictor is (e.g., Pulakos et al., 2002), these efforts tend to differ in their conceptualizations. Similarly, it may be that “adaptability” does not represent a distinguishable attribute per se (like GMA), but rather an enduring characteristic of an attribute(s). For example, in personality psychology, there is the concept of trait strength or “traitedness,” which suggests that for each individual specific traits are more dominant than others and will consistently (and significantly) influence one’s actions across a wide range of situations. Comparable research by Mischel and colleagues (Mischel, 2004; Mischel & Shoda, 1998) has found that individuals’ personality, as described by self and others, can systematically differ across situations, such that people may be consistently extraverted in certain situational contexts (e.g., work) but introverted in others (e.g., home). One interpretation of these findings is that “adaptability” is not so much an attribute as it is an enduring feature (or characteristic) of attributes. Comparable issues have been raised with attributes, such as problem-solving and critical thinking (e.g., Campbell, 2003), indicating that these attributes are *domain*-specific. While these issues raise practical concerns, we do not believe they should automatically rule-out these attributes from consideration.

Finally, it should be noted that many Air Force stakeholders will likely agree with the importance of many of these attributes for officer selection – indeed, as Tables 15 and 16 illustrate, several of these attributes are potentially being measured by current selection practices. The problem, as discussed previously, is that (a) what attributes are being assessed has not been clearly articulated, (b) the degree to which current practices provide accurate and valid assessments of one or more of these attributes have not been established, and (c) their linkages to officer performance (competency) requirements have not been well specified.

Step 2: Identifying and Evaluating Assessment Methods

What We Did

Although it has not always been the case operationally, all of the predictor attributes identified in the preceding section can be assessed in multiple ways. To keep our evaluation focused, we identified five standard assessment methods by which one or more of these attributes have been (or could be) measured. They were (a) standardized tests, (b) interviews, (c) self-reports, (d) ratings, and (e) simulations. Generally, these methods were selected because they (a) potentially represented the most feasible and commonly used assessment methods, and (b) have frequently been the focus of comparable military-sponsored research efforts to expand the predictor space (cf. Campbell & Knapp, 2001; Knapp et al., 2004). The major types and features commonly associated with each method are briefly summarized in Table 17.

Building off of the effectiveness criteria formulated earlier in the report, we focused on the following considerations in our evaluation:

- Coverage of critical predictor attributes and efficiency for predicting Air Force competency (performance) requirements;
- Expected operational validity, as demonstrated (or suggested) by past research and/or experience with similar methods;
- Potential for adverse impact;

- Development and administration costs;
- Potential for stakeholder ―buy in” and credibility;
- Potential for computer- or Internet-based administration;
- Suitability for use in a military assessment context, specifically for officer selection and assessment.

Using these criteria, we conducted a comprehensive review of information covering one or more of these assessment methods. Relevant sources included meta-analyses and individual research studies (e.g., McDaniel et al., 2001; Nguyen, McDaniel, & Whetzel, 2005; Schmidt & Hunter, 1998), narrative reviews (e.g., Borman et al., 2003; Guion, 1998; Hough, Oswald, & Ployhart, 2001; Sackett, Berry, & Wiemann, 2005), technical reports, conference papers, and workshops documenting predictor assessment development efforts, specifically those conducted for the military (e.g., Bowden, Keenan, Knapp, & Heffner, 2004; Campbell & Knapp, 2001; Knapp et al., 2004; Michael, Chen, Janega, Farmer, Eller, & Nayak, 2005), and information from test publishers (e.g., The College Board). In addition, per our earlier recommendation about replacing AFOQT-AA subtests (Verbal Analogies, Arithmetic Reasoning, Word Knowledge, and Math Knowledge) with the SAT Reasoning Test, we compared the SAT and several alternative standardized tests of GMA and other cognitive attributes (e.g., ASVAB, GRE, LSAT, GMAT), which represent the most potentially feasible substitutes (i.e., those with an existing infrastructure to support and sustain long-term development and large-scale administration), against the same criteria.

Findings

Table 18 provides a comparative summary of the different assessment methods by the evaluation criteria. No method is inherently ―better” than another, as each has its relative advantages and disadvantages. The following substantive points should be taken from this review:

First, those methods offering the greatest amount of coverage (i.e., number of attributes assessed) relative to their costs are roughly (in descending order of expected operational validity) (a) situational judgment tests (SJTs), (b) (situational) interviews, (c) ratings, (d) personality-temperament inventories, (e) biodata-experience inventories, and (f) (behavioral description) interviews. In particular, SJTs are attractive as they (a) are reasonably cost-effective, (b) produce, on average, operational validities comparable to that of standardized aptitude-ability tests, (c) have been found to predict performance across military and civilian settings, (d) produce less adverse impact, on average, than standardized ability-aptitude tests, and (e) provide an approach for assessing cross-functional skills (e.g., social skills) without some of the same drawbacks as self-reports (Borman et al., 2003; McDaniel et al., 2001; Nguyen et al., 2005; Pearlman, 2003). In addition, SJTs can provide a realistic job preview useful for minimizing non-academic attrition by presenting prospective candidates – specifically civilians with no prior exposure to the military – with situations comparable to those they will encounter during commissioning (or on-the-job). As with interviews, the major disadvantage of SJTs has been establishing what they are measuring, but most likely this is because the items are multi-dimensional in nature.

Table 17. Summary of Methods for Assessing Leader Attributes

Assessment Method	Type	Description	
Standardized Tests	Traditional Multiple-Choice	• Test-takers respond to standardized set of items, choosing one correct (or best) response from 3-to-5 options.	• Presentation of items can be linear or nonlinear (e.g., CAT). • Objectively scored.
	Performance-Oriented Multiple-Choice	• Test-takers respond to standardized set of items, using different response formats (e.g., matching, drag-and-drop). • Visual aids to reduce reading and enhance realism (e.g., photos, figures).	• Presentation of items can be linear or nonlinear (e.g., CAT). • Animation to enhance realism. • Objectively scored.
	Constructed Response	• Test-takers construct response (e.g., short answer, paragraph, essay, and so on) to standardized set of items.	• Can be scored objectively or by expert judgment.
Interviews	Behavioral Description	• Based on premise that best predictor of future behavior is past behavior. • Interviewees are asked to provide past behavioral exemplars to similar performance-relevant situations (e.g., ―Tell me about a time when you…”).	• Can differ in level of structure (or standardization). • Responses typically scored by interviewer (or panel of interviewers).
	Situational	• Interviewees respond to hypothetical performance-relevant scenarios (or critical incidents).	• Can differ in level of structure (or standardization). • Responses typically scored by interviewer (or panel of interviewers).
	Fact-Based	• Interviewees respond to objective questions measuring role- or job-related knowledge.	
Self-Reports	Personality-Temperament Inventories	• Test-takers respond to statements (or other test stimuli) based on degree to which it is self-descriptive (or not).	• Can be checklists, scaled-responses, or multiple-choice (or force-choice) format.
	Biodata-Experience Inventories	• Based on premise that best predictor of future behavior is past behavior. • Measures past experiences, training (or certifications), accomplishments, and so on.	• Can be fact- (overt) or personality-based (covert). • Scoring based on rational or empirical keying.

Table 17. Summary of Methods for Assessing Leader Attributes (cont'd)

Assessment Method	Type	Description	
Ratings	Performance-Potential Ratings	• Subjective appraisals of candidates' performance and/or potential on one or more role- or job-relevant dimensions.	• Can be completed by candidate's supervisor, peers, direct reports, and/or others (e.g., customers). • Can be rankings (relative or absolute), checklists, or anchored ratings scales.
Simulations	Situational Judgment Tests (SJTs)	• Performance-relevant problem scenarios depicted in writing or through video. • Test-takers evaluate various possible actions (i.e., what s/he ―should-do" or ―would-do").	• Focus is typically on judgment rather than knowledge per se. • Can be multiple-choice or constructed response. • Scoring key based on expert judgment.
	Path (Directed) Simulation	• Test-takers are presented with a computer simulation of a problem scenario.	• Test-takers progress through the simulation, stopping at various points to answer questions.
	Open (Interactive) Simulation	• Test-takers are presented with a computer simulation of a problem scenario.	• Test-takers progress through the simulation, stopping at various points to answer questions.

Table 18. Overview Comparison of Different Assessment Methods

Method	**Leader Attribute Coverage**	**Operational Validity**	**Adverse Impact**	**Operational Costs**		**Stakeholder “Buy In”**		**Technology Enabled**	**Military Suitability**
				Develop	Admin	Candidates	Organization	Y/N	Y/N
Standardized Tests									
Traditional MC	Low	Low to High	Med to High	Med to High	Low	Med	High	Y	Y
Performance-Oriented MC	Low	Low to High	Med	Med to High	Low to Med	High	High	Y	Y
Constructed Response	Low	Low to High	Med	Med to High	Med	Med to High	High	Y	Y
Interviews									
Behavioral Description	High	Low to Med	Low	Med	Med to High	High	Med to High	N	Y
Situational	High	Med to High	Low	Med	Med to High	High	Med to High	N	Y
Self-Reports									
Personality-Temperament Inventories	High	Low to Med	Low	Med	Low	Low	Low to Med	Y	Y
Biodata-Experience Inventories	High	Low to Med	Low	Med	Low	Low to Med	Low to Med	Y	Y
Ratings									
Performance-Potential	High	Low to Med	Low to Med	Med	Low	High	High	Y	Y

Simulations									
Situational Judgment Tests (SJTs)	Med to High	Med	Low to Med	Med to High	Low	High	High	Y	Y
Path-Open Simulations	Med to High	Med to High	Low to Med	High	Med to High	High	High	Y	Y

Notes. Operational validities categorized using Cohen’s (1988) effect size classifications for correlations (*Low* = .10; *Med* = .30; *High* = .50). Adverse impact similarly categorized using Cohen’s (1988) effect size classifications for standardized mean differences (d values) (*Low* = .20; *Med* = .50; *High* = .80).

Second, methods with potentially greater fidelity (i.e., where responding to the test more closely elicits the processes being assessed) and thereby greater operational validity, such as path-open simulations, are prohibitively more costly, which accounts for the infrequency of their use. Related to this, such assessments frequently require computer- or Internet-based administration. Nevertheless, computer- or Internet-based assessment can be advantageous because it (a) avoids many of the security hazards and costs associated with a large-scale paper-based assessment program (including administration costs associated with updating and redeploying new forms), (b) are frequently found more appealing and credible by test-takers than paper-based assessments of the same attributes, and (c) can facilitate the delivery of immediate and fairly comprehensive feedback to test-takers and decision-makers (Jones & Dages, 2003). The Army and Navy have investigated and prototyped large-scale assessment programs, specifically for enlisted promotions, that make use of higher fidelity assessments (cf. Baisden, Schultz, & Lewis-Brown, 2004; Knapp & Campbell, 2005). DMDC is pursuing initiatives related to an Internet-based version of the ASVAB that could similarly make higher fidelity assessments more feasible. While creating such a system ―from scratch” may be prohibitive, taking advantage of the infrastructure and lessons learned created by these efforts holds promise for opening-up or enhancing the assessment of leader attributes that are difficult to measure using paper-and-pencil forms. This could be especially true for assessing some of the higher-level cognitive aptitudes-abilities and cross-functional skills (e.g., social skills), which are more procedurally- or process-based and thereby have historically proven difficult to assess.

Third, as can be seen from Table 19, all other things being equal, the SAT Reasoning Test is arguably the most preferable candidate for replacing the AFOQT-AA subtests. Generally speaking, with the exception of the ASVAB, all of the other alternatives (GRE, LSAT, GMAT) assess higher-level cognitive aptitudes-skills and cross-function skills (e.g., critical thinking skills, verbal-written communication skills) not comprehensively assessed by the AFOQT, which could increment the prediction of officer success. Of these alternatives, the one that is currently in operational use by services for making junior officer selection and assessment decisions, and which has received the most extensive study equating it with the AFOQT, is the SAT. Additionally, the SAT is more likely to enjoy greater credibility (―face validity”) among Air Force stakeholders, as it is not specifically geared towards specialized educational (or academic) training, such as the LSAT (law school) or GMAT (business school). Having said that, and as discussed previously, the SAT Reasoning Test has several drawbacks (a) it is not DoD-owned, although the population and institutional goals the SAT serves is comparable to those of the services, and (b) since its scores could be at least several years old at time of application some subtest scores (e.g., writing) might not provide a completely up-to-date and accurate assessment of candidates’ current skill level (e.g., due to maturation and training), although this will primarily be an issue for those applying to OTS. The SAT Reasoning Test emerges as the best and most practical alternative relative to other readily accessible, large-scale standardized tests.

Table 19. Overview Comparison of Readily Accessible Standardized Tests of GMA

Criteria	**Test**				
	SAT	*ASVAB*	*GRE*	*LSAT*	*GMAT*
Leader Attribute Coverage	Measures critical thinking, general reasoning , verbal-written communication, and quantitative ability .	Measures GMA (AFQT), along with specific aptitudes and job family-specific knowledge and skills (e.g., Electronics).	Measures critical thinking , general reasoning, verbal-written communication ,and quantitative ability.	Measures critical thinking, general-analytical reasoning, and verbal-written communication.	Measures critical thinking, general-analytical reasoning, problem-solving, verbal-written communication, and quantitative ability.
Operational Validity	Earlier versions produce comparable operational validities for commissioning and post-commissioning criteria as AFOQT-AA-V-Q. New version potentially higher than AFOQT-AA-V-Q, as test includes assessments of higher-level cognitive aptitudes.	Generally, operational validities comparable to, and in some cases somewhat higher than, AFOQT for technical training and job performance criteria for enlisted personnel and NCOs.	Potentially higher than AFOQT-AA, as it includes assessment of higher-level cognitive skills.	Potentially higher than AFOQT-AA, as it includes assessment of higher-level cognitive skills, but no quantitative ability.	Potentially higher than AFOQT-AA, as it includes assessment of higher-level cognitive skills.
Operational Costs	No development and/or administration costs.	No development costs. Could be administration costs depending on agreement with DoD.	No development costs. Administration costs to applicant for taking test (e.g., $115-$140 registration fee).	No development costs. Administration costs to applicant for taking test (e.g., $115 registration fee).	No development costs. Administration costs to applicant for taking test (e.g., $250 registration fee).

Note. SAT = SAT Reasoning Test; ASVAB = Armed Services Vocational Aptitude Battery; GRE = Graduate Record Examination; LSAT = Law School Admission Test; GMAT = Graduate Management Admission Test.

Table 19. Overview Comparison of Readily Accessible Standardized Tests of GMA (cont'd)

Criteria	**Test**				
	SAT	*ASVAB*	*GRE*	*LSAT*	*GMAT*
Stakeholder "Buy In"	Expect to be high. Currently used operationally for officer (and scholarship) selection by Air Force decision-makers.	Potentially high, though expect concerns about possible differences in test difficulty and efficiency (e.g., from completing longer test).	Expected to be comparable to SAT. Candidates not likely to respond positively to registration costs without Air Force reimbursement.	Potentially low, as test is specifically geared towards entry to specialized training (law school)l.	Could be mixed, as test is specifically geared towards entry to specialized training (business school).
Technology Enabled	Paper-and-pencil administered.	CAT-administered version (CAT-ASVAB); DoD exploring Internet-administered version (iCAT-ASVAB).	CAT-administered version.	Paper-and-pencil administered.	CAT-administered version.
Military Suitability	Used operationally by AFROTC in place of AFOQT-AA when selecting for POC. Used by other services (e.g., Army, Navy) for officer selection (civilians).	General Technical composite (ASVAB) used by Army for selection to Officer Candidate School and for selecting enlisted personnel into Army ROTC.	Currently not used by any of the services for (junior) officer selection.	Currently not used by any of the services for (junior) officer selection.	Currently not used by any of the services for (junior) officer selection.
Other Considerations		DoD/DMDC investigating and piloting new subtests to assess basic computer skills and written communication skills.	ETS plans to implement changes in Oct 06 that place greater emphasis on higher cognitive skills, assess quantitative skills more consistent with those required in grad school, and more quantitative questions involving "real life" scenarios. Can re-take test up to 4 times within 12-month period.	Administered only 4 times per year.	Can re-take test up to 4 times within 12-month period.

Step 3: Recommendations for a 21st Century Officer Assessment “Toolkit”

From the preceding findings, we constructed an officer assessment ―toolkit” that could be used operationally to meet the Air Force’s needs and performance imperatives. Consistent with the Air Force’s goals for selection system success, and following earlier recommendations, the ―toolkit” primarily aimed to:

- Maximize coverage of critical predictor attributes, so as to optimize the prediction of officer success (commissioning and post-commissioning);
- Ensure a productive balance between new and existing (or enhanced versions of existing) measures to retain, and promote, stakeholders’ credibility and acceptance of the proposed ―toolkit”;
- Minimize adverse impact to promote Force diversity;
- Incorporate components that compliment and support other HR systems, specifically recruitment and training;
- Promote a reasonable level of standardization across the different commissioning sources in the predictor attributes assessed (and how they are assessed);
- Identify and leverage existing operational (or experimental) measures available within the Services, some of which have computer or Internet-enabled versions, to maximize the Air Force’s ROI and to facilitate standardization across the Total Force;
- Be reasonably affordable to develop, implement, and sustain.

Table 19 summarizes the ―toolkit’s” main features and components, spanning recruitment to commissioning. As the table illustrates, the ―toolkit” incorporates and structures assessments meant to compliment and substantively support other HR systems, mainly recruitment and training. The following describes these features by the HR function they intend to serve.

Officer Recruitment and Commissioning (Training)

On the recruitment side, these assessments aim to provide recruiters and/or prospective candidates with diagnostic information they can use to measure their readiness and prospects for commissioning. For example, for prospective candidates, these tools can be employed to (a) self-assess their potential strengths and weaknesses as candidates, and identify what they need to do to improve their chances for selection; (b) explore their potential ―fit” with the Air Force and the role of officer, and possibly their ―fit” with specific specialties, such as flying (pilot-navigator); and (c) their expectations and beliefs about the Air Force and military service. Chief among the advantages of these assessments are that they (a) facilitate the ―selecting out” of less qualified candidates, (b) supply highly motivated candidates with fairly specific, targeted feedback and recommendations on areas for improvement, and (c) assist in the identification of qualified candidates with an interest in hard-to-fill specialties. Taken together, these assessments function to improve the quality of the officer applicant pool and promote a favorable image of the Air Force as an organization invested in its people.

Table 19. Proposed Officer Assessment “Toolkit”

Commissioning Source	***HR Function***			
	Recruitment	**Selection**		**Commissioning-Training**
		Officer Hurdle	***Technical Hurdle***	
AFROTC	• Person-Environment (P-E) ―Fit” Assessment* • SAT Reasoning Test (Critical Reading, Writing, Math)* • AFQT (ASVAB)* • Sample, non-operational items from AFOQT-EX* • Experience and activities record*	• Relative Standing Score-EX (RSS-EX) • Cumulative GPA • PFT • SAT Reasoning Test (Critical Reading, Writing, Math) • AFOQT-EX	• Pilot-Navigators (e.g., Pilot-Nav Composites; TBAS) and so on…	• Person-Environment (P-E) ―Fit” Assessment* • AFOQT-EX* • Experience and activities record* • Pilot-Navigators (e.g., Pilot-Nav Composites; TBAS)*
OTS	• Person-Environment (P-E) ―Fit” Assessment* • SAT Reasoning Test (Critical Reading, Writing, Math)* • AFQT (ASVAB)* • Sample, non-operational items from AFOQT-EX* • Experience and activities record*	• Recruiter/Interviewer Evaluation-EX • GPA • PFT • SAT Reasoning Test (Critical Reading, Writing, Math) • AFOQT-EX	• Pilot-Navigators (e.g., Pilot-Nav Composites; TBAS) • Academic Major GPA and so on…	• Person-Environment (P-E) ―Fit” Assessment* • AFOQT-EX* • Experience and activities record* • Pilot-Navigators (e.g., Pilot-Nav Composites; TBAS)*

Note. EX = enhanced version. * = Intended for diagnostic feedback and not for operational decision-making.

While not an exhaustive list, the following tools were identified specifically for this purpose, some of which are currently in use, with examples of appropriate existing and/or experimental measures developed and owned by one or more of the Services where applicable:[28]

- Person-environment –fit” measures, specifically those assessing prospective candidate’s fit with a leadership role and with Air Force culture and their expectations and beliefs about the Air Force and military service. Example measures include: JOIN (Jobs & Occupational Interests in the Navy); AVOICE (Army Vocational Interest Career Examination); Interest Finder Questionnaire (IFQ-DMDC); Work Preferences Survey (WPS-Army); and Army Beliefs Survey (ABS).
- SAT Reasoning Test – both total and subtest (Critical Reading, Writing, Math) scores.
- AFQT – for currently enlisted personnel considering pursuing an officer commission.
- Sample items from an enhanced version of the AFOQT (AFOQT-EX), which serves to provide a preview of AFOQT and an initial self-assessment of how a candidate might do.
- An experience and activities record, a checklist-type self-report of how frequently prospective candidates’ have engaged in leadership-related activities and other experiences (work-related and so on) relevant to their potential and preparation for commissioning—and possibly for specific specialties. We know of no example measures directly applicable to an officer candidate population, though there are several geared towards officer promotions [e.g., Experiences and Activities Record (ExAct-Army)].

On the commissioning side of the equation, several of these measures could be employed to (a) facilitate –selecting out” from commissioning of less motivated officer candidates, and (b) assist candidates in identifying post-commissioning specialty assignments, mainly for AFROTC. Similarly, scores obtained during the selection process (e.g., AFOQT), when supplement with productive feedback, could be used by candidates to make the most out of their commissioning program (e.g., by identifying specific training and development needs). AFROTC and OTS could also use these scores, in the aggregate (i.e., for a particular cohort of candidates), in planning, designing, and implementing their respective commissioning programs.

Officer Selection

As recommended earlier, selection for officer potential is meaningfully separated from selection for technical specialty potential. How this is accomplished could take one of many forms depending on what works best for the Air Force and the respective commissioning source. As for assessment content, the proposed –toolkit” generally retains many of the same basic measures of the existing systems, particularly those (e.g., RSS, Recruiter/Interviewer Evaluations) that currently play a significant role in officer selection. Where they substantively differ is as follows:

---

[28] In several instances, computer-based versions of these measures have already been developed.

- As recommended previously, use of the SAT Reasoning Test (Total, Critical Reading, Writing, Math) as an assessment of GMA. For enlisted personnel where SAT scores are otherwise unavailable or more than a certain number of years old, the AFQT, which is also a reliable measure of GMA (e.g., Campbell & Knapp, 2001; Sperl et al., 1992; Welsh et al., 1990), could potentially be used in its place.
- Use of enhanced versions of the Relative Standing Score (RSS) and Recruiter/Interviewer Evaluations where ―officer potential” has been substantively specified in line with Air Force competency (performance) requirements and some needed standardization (structure) has been introduced into assessment to ensure consistency within and across commissioning sources.
- Use of an enhanced AFOQT that (a) emphasizes higher-order cognitively-oriented aptitudes predictive of officer success (commissioning and post-commissioning) not covered by GMA, and (b) supplies standardized assessments of cross-functional skills and personal qualities (e.g., personality) to supplement and to validate those obtained from the RSS and recruiter/interviewer evaluations. While the exact specification of this enhanced AFOQT requires additional analyses and decisions on part of Air Force stakeholders, Table 20 offers an initial take. As evident from the table, the only new prescribed addition is a situational judgment test (SJT) targeting cross-functional skills, such as social skills and problem-solving. Recently, there have been several efforts by the Services to construct SJTs from which the Air Force could capitalize in constructing this measure (cf. Bowden et al., 2004; Knapp et al., 2004). Beyond this, the remainder of the test content is flexible. This space could be used to (a) incorporate existing subtests (e.g., Table Reading for measuring Information Processing Skills) assessing one or more targeted attributes not sufficiently covered by the SAT Reasoning Test, (b) include shorter ―marker” tests comparable to the SAT subtests that could be used to supply a more up-to-date, or more military-specific, assessment of select attributes measured by these subtests, or (c) include subtests mainly intended for diagnostic and/or other assessment purposes besides officer selection. Either way, the proposed enhanced AFOQT is sensitive towards the practical time constraints related to its administration, as it seeks to stay within the same total timeframe (3.5 hours) as the current AFOQT (Form S).

Table 20. Proposed AFOQT-EX

Subtest	**Number of Items**	**Testing Time (in Minutes)**
Experimental A	--	--
Experimental B	--	--
Experimental C	--	--
Table Reading (TR)	40	9
Situational Judgment Test (SJT)	35-40	45-50
Self-Description Inventory (SDI+)	220	40
Totals	**300**	**99**

Implications of Implementing the Proposed “Toolkit”

As discussed earlier, the proposed assessment ―toolkit” seeks to address several of the Air Force’s goals for selection system success, while incorporating many of our earlier recommendations. As Table 21 illustrates, the proposed measures comprising the ―toolkit” provide Air Force decision-makers with reasonable coverage of the critical leader attributes previously identified. Additionally, the number of new measures needing to be developed ―from scratch” is generally limited since most of the measures proposed represent enhanced versions of current ones (e.g., RSS) or require modifications to existing tools developed for or by the other services. Nevertheless, implementing the proposed ―toolkit” will require:

- Re-evaluating retesting policies, specifically as they apply to the SJT and similar assessments (e.g., SDI+);
- Revisions to test preparation guides and potentially substantive changes in how data collected from new assessments is used, distributed, and disseminated (e.g., what kinds of feedback prospective candidates receive and in what form);
- Publication and/or production of new assessment tools, although structurally the proposed ―toolkit” does not call for substantive changes in how these processes are currently organized;
- The setting of norms, cut scores, and/or other comparable aids for new assessments to assist decision-makers, and research to inform the construction of these aids;
- Modifications (or additions) to the AFOQT composites currently used in officer selection—ideally in line with the Air Force’s competency (performance) model but sufficiently credible and intuitive to operational decision-makers;[29]

---

[29] As it relates to the SAT these changes are not expected to be substantial since AFOQT-AA, AFOQT-V, and AFOQT-Q (conceptually) map onto related SAT subtest scores. Additionally, this is already being done operationally by AFROTC. Generally speaking, the most significant change to composites will come from adding assessments targeting cross-functional skills and personal qualities (e.g., personality) not covered by the SAT (or AFOQT). At some point, however, this will need to be done for the SDI+ to go operational.

Table 21. Coverage of Recommended Predictor Attributes by Proposed Measures

	Proposed Measures										
							AFOQT-EX				
Recommended Predictor Attributes	P-E Fit	Exp Act	RSS-EX	Recruit Eval-EX	SAT	ASVAB (Opt)	SDI+	SJT	TR	Exp Test A	Exp Test B
Cognitive-Integrative Complexity										??	??
Creative-Divergent Thinking (Skills)										??	??
Critical Thinking (Skills)					X					??	??
General-Analytical Reasoning (Skills)					X					??	??
Information Processing (Skills)						X			X		
Problem-Solving (Skills)								X		??	??
General Mental Ability					X	X					
Verbal-Written Communication Skills		X		X	X	X					
Achievement Orientation		X	X	X							
Adaptability		X	X	X				X			
Conscientiousness-Integrity		X	X	X			X	X			
Core Self-Evaluation (Emotional Stability)		X	X	X			X				
Energy Level-Potency		X	X	X			X				
Motivation to Lead	X	X	X	X			X	X			
Motivation to Learn		X	X	X			X				
Openness to Experience-Cultural Tolerance		X	X	X			X	X			
Service-Team Orientation	X		X	X			X				
Self Directed Learning & Development Skills		X	X	X							
Self-Management Skills		X	X	X				X			
Social-Interpersonal Skills		X	X	X				X			

Note. X = Designed to measure.

- As with the current AFOQT, research to develop and validate new assessments, specifically those used for officer selection. Additionally, long-term investment to periodically replace forms (as needed), although taking advantage of recent and future efforts by other services to construct comparable assessments would help minimize these costs;
- As needed, additional research demonstrating the practical impact (e.g., who is selected, capability to meet end-strength requirements, etc.) and predictive efficacy of alternatives, such as the new SAT Reasoning Test, to assure key Air Force stakeholders of the value-added and absence of adverse effects on component needs and performance requirements;
- To facilitate implementation, effective communication (e.g., via training, policy guidance, and/or other documentation) of changes to high-level AFROTC and OTS decision-makers on down to MEPs test administrators (e.g., what new assessments mean; how to use them, and so on);
- A programmatic system, comparable to the CEMU, for maintaining and sustaining these changes;
- Most importantly, the committed support and partnerships among relevant Air Force stakeholders to facilitate implementation and maximize ROI.

Moving Ahead: A Summary of What is Needed and Next Steps

What is Needed

For the Air Force and other services, officer selection has and continues to carry significant strategic importance. Because the Air Force's junior officers represent its future, their selection constitutes one of the most critical and substantial investments it makes. Without the right investment, the Air Force cannot ensure that officer selection continues to meet its needs and performance imperatives. Consistent with this, this report aimed to systematically review current selection systems and their components and to evaluate possible alternatives that could significantly enhance officer quality. From this effort, a number of actionable recommendations for possible future enhancements were generated. Table 22 summarizes those recommendations that emerged as most critical.

Next Steps

An Implementation Road Map

Figure 8 presents a road map, along with critical decision points, for successfully implementing the recommendations summarized in Table 22. As Figure 8 illustrates, we envision four steps (or phases) to implement these recommendations (e.g., Gallagher, Joseph, & Park, 2002).

Table 22. Recommendations of What is Needed in Junior Officer Selection

Selection System Focus	**Time Frame**	
	Short-Term	*Long-Term*
Strategy	• Develop and implement a competency (performance) model and align officer selection systems accordingly, so that performance requirements – to include officer retention – can (and do) measurably drive selection. • Implement a sufficiently comprehensive criterion assessment system that enables the measurement of desired results (e.g., officer performance and retention), so that the success of officer selection systems and their components can be effectively assessed. • Investigate and take advantage of ways in which current (or proposed) selection tools could be integrated with recruitment and training to compliment and enhance the Air Force's ROI.	• Ensure there is a formal, programmatic system in place for sustaining and maintaining linkages between officer selection systems (and their components) and future changes to competency (performance) requirements and other critical imperatives (e.g., retention).
Content	• Partner with AFRS and AFROTC (and USAFA) to standardize the specification (and conceptualization) of predictor attributes in use for officer selection. • Maximize coverage of predictor attributes not currently assessed (e.g., cross-functional skills), whose addition could significantly increment the prediction of officer success and promote Force diversity.	• None.
Method, Design, & Implementation	• Meaningfully and measurably separate assessment of officership from the assessment of technical expertise. • Partner with AFRS and AFROTC to structure existing officer selection measures to promote and facilitate standardization across commissioning sources. • Partner with AFRS to systematically review and evaluate specific assessments (e.g., letters of recommendation) whose ROI is potentially limited. • Investigate implementation issues and related concerns (e.g., faking and coaching) that could preclude the use of the SDI+ as an officer selection & classification tool.	• Systematically study and investigate the possibility of using the new SAT Reasoning Test in place of the GMA-specific subtests in current AFOQT. • Investigate possibilities for a joint-service (junior) officer selection battery.

Figure 8. Road Map for Implementing Recommendations

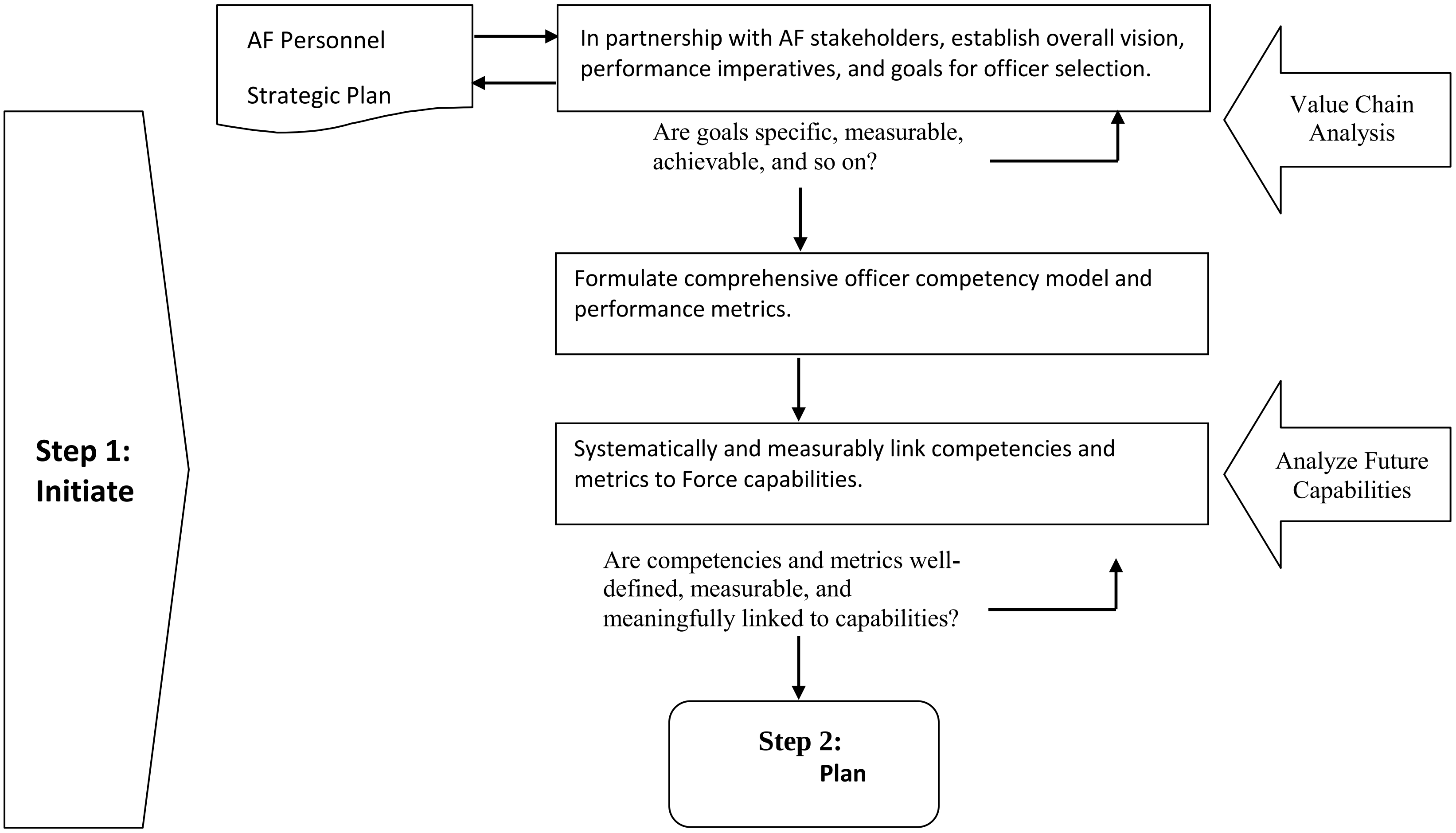

Step 1:
Initiate
AF Personnel
Strategic Plan
In partnership with AF stakeholders, establish overall vision, performance imperatives, and goals for officer selection.
Value Chain Analysis
Are goals specific, measurable, achievable, and so on?
Formulate comprehensive officer competency model and performance metrics.
Systematically and measurably link competencies and metrics to Force capabilities.
Analyze Future Capabilities
Are competencies and metrics well-defined, measurable, and meaningfully linked to capabilities?
Step 2:
Plan

Figure 8. Road Map for Implementing Recommendations (cont'd)

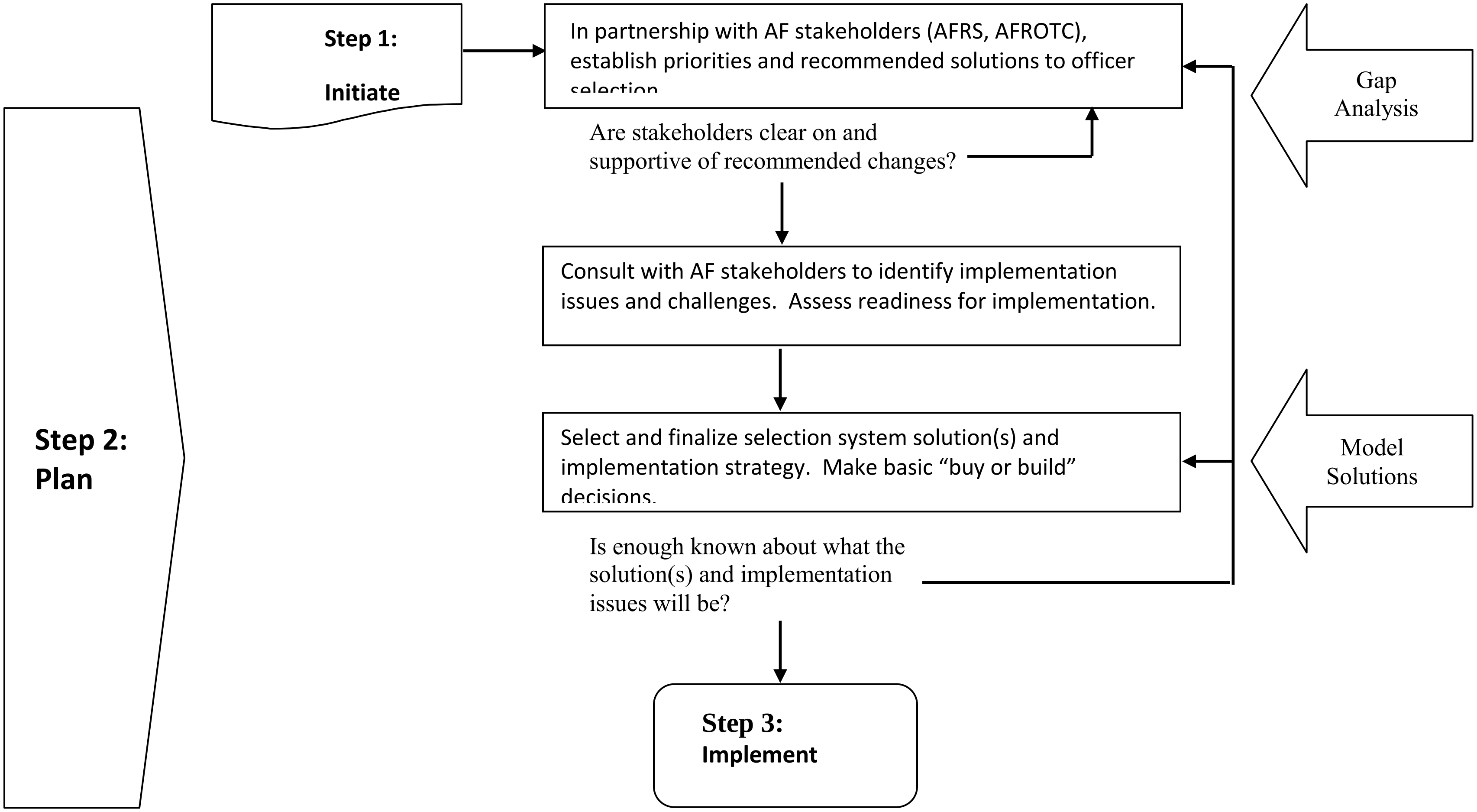

Step 1:
Initiate
In partnership with AF stakeholders (AFRS, AFROTC), establish priorities and recommended solutions to officer selection
Gap Analysis
Are stakeholders clear on and supportive of recommended changes?
Consult with AF stakeholders to identify implementation issues and challenges. Assess readiness for implementation.
Step 2: Plan
Select and finalize selection system solution(s) and implementation strategy. Make basic "buy or build" decisions.
Model Solutions
Is enough known about what the solution(s) and implementation issues will be?
Step 3: Implement

Figure 8. Road Map for Implementing Recommendations (cont'd)

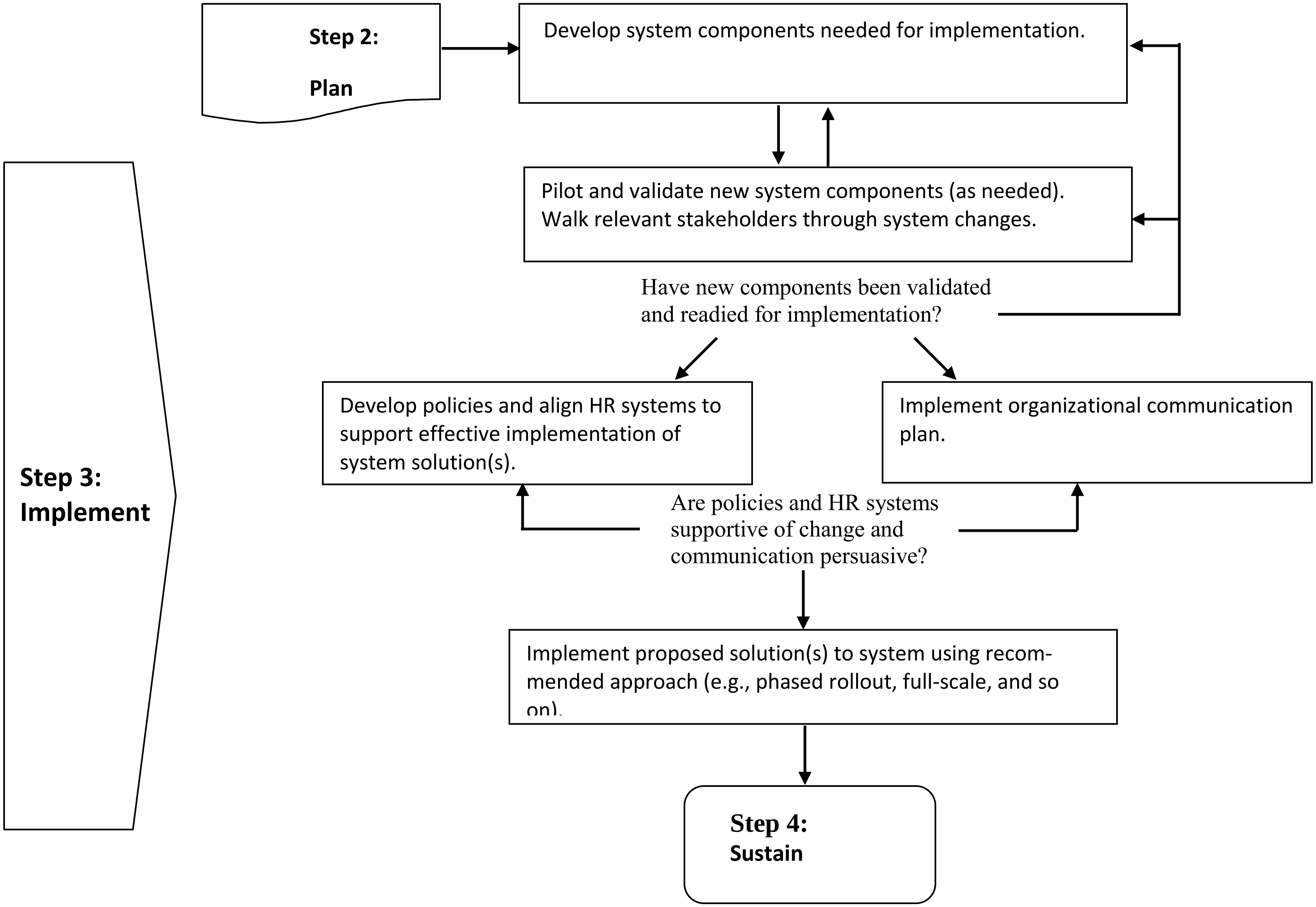
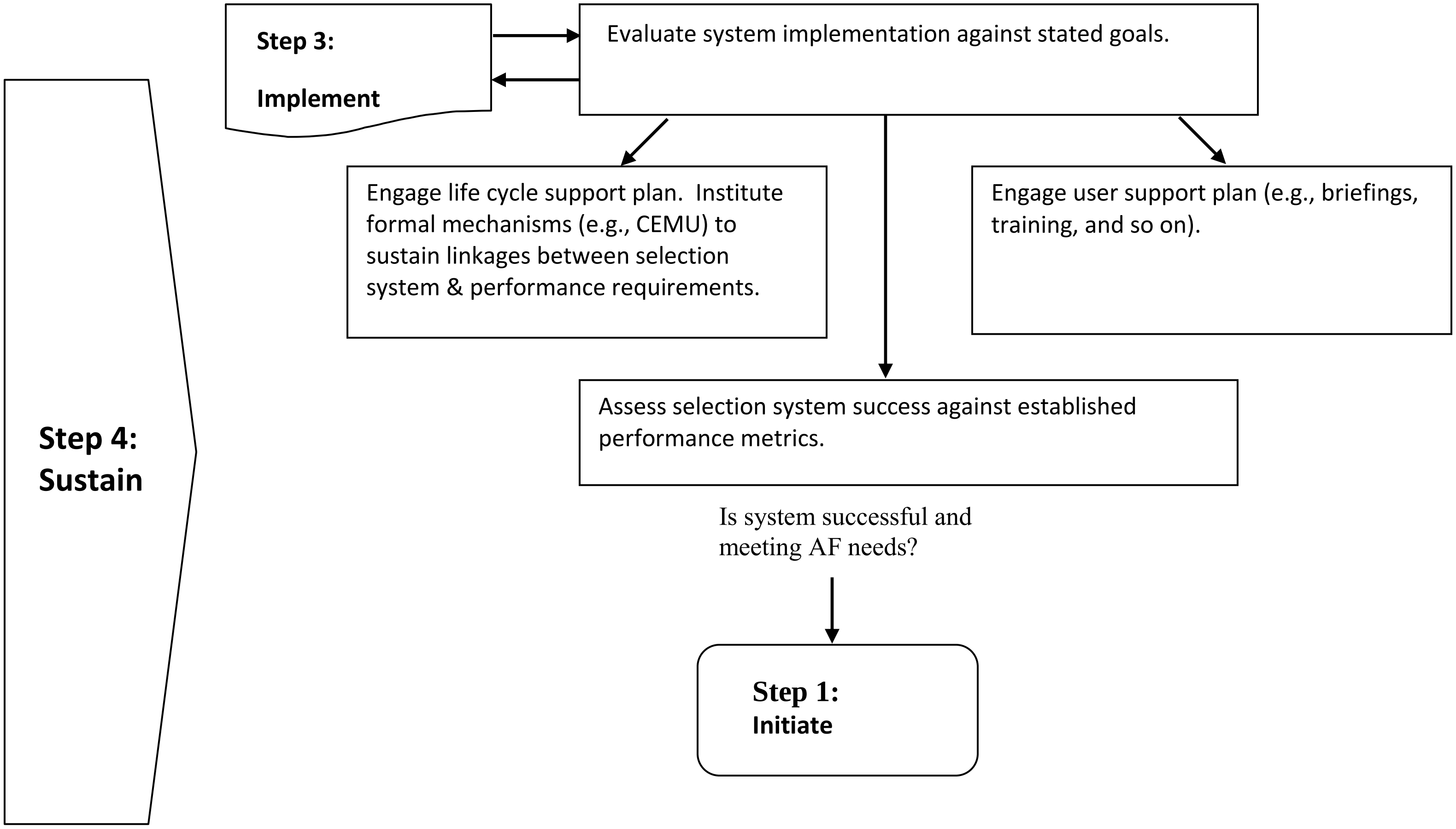

Step 2:
Plan
Develop system components needed for implementation.
Pilot and validate new system components (as needed).
Walk relevant stakeholders through system changes.
Have new components been validated and readied for implementation?
Step 3:
Implement
Develop policies and align HR systems to support effective implementation of system solution(s).
Implement organizational communication plan.
Are policies and HR systems supportive of change and communication persuasive?
Implement proposed solution(s) to system using recommended approach (e.g., phased rollout, full-scale, and so on).
Step 4:
Sustain

Figure 8. **Road Map for Implementing Recommendations (cont'd)**


Step 3:
Implement
Evaluate system implementation against stated goals.
Step 4:
Sustain
Engage life cycle support plan. Institute formal mechanisms (e.g., CEMU) to sustain linkages between selection system & performance requirements.
Engage user support plan (e.g., briefings, training, and so on).
Assess selection system success against established performance metrics.
Is system successful and meeting AF needs?
Step 1:
Initiate

Step 1: Initiate. This step consists of enacting the recommendations under ―Strategy.” Building off of the *Personnel Strategic Plan*, it begins with articulating the vision, performance imperatives, and goals driving officer selection and its intersection with other HR functions and processes (e.g., recruitment, training, and so on). Table 9 summarizes our attempt to formulate these imperatives and goals, although additional work is expected. To ensure that goals are reasonably specific and measurable, value chain analysis and other analytic tools could prove useful, especially for identifying and analyzing the linkages between selection and other critical HR systems. Overall, successfully completing this step should result in (a) securing the commitment and support of relevant Air Force stakeholders, at all levels, to making changes to enhance the current officer selection systems, (b) a comprehensive officer competency model, and (c) performance metrics, includes diversity-based metrics, useful for assessing selection system success.[30]

Step 2: Plan. The primary objectives of this step are to select which proposed solution(s) and changes to implement and to plan for their implementation. While admittedly not exhaustive, the recommendations under ―Content” and ―Method, Design, & Implementation” outline a number of potential solutions and changes that are expected to enhance officer selection. Conducting a gap analysis (and/or the research recommended below) should facilitate the prioritizing of these recommendations and filling in their specifics (e.g., *how* much to operationally weight officership and technical expertise and using what selection model). Because of their responsibility for managing and administering the officer selection process, and consistent with these recommendations, partnering and consulting with relevant Air Force stakeholders (e.g., AFRS, AFROTC) is critical. Findings from the aforementioned gap analysis and other research should be instrumental in securing the commitment and support of key stakeholders. When completed, this step should produce (a) the commitment and support of relevant stakeholders, specifically AFRS and AFROTC, to implementing proposed solution(s), (b) a clear specification of what solution(s) and changes to enact, and (c) a strategy (i.e., timeline with critical milestones, specification of roles and responsibilities, an evaluation plan, and so on) for implementing them.

Step 3: Implement. In accordance with the implementation strategy crafted in Step 2, this step involves developing and deploying specified solutions and changes to officer selection. As with the earlier steps, relevant stakeholders' commitment and involvement to process is critical. As can be seen from the road map, and consistent with recommended practices for managing organizational change (cf. Gallagher et al., 2002), successfully fostering and maintaining this commitment will require (a) policies and HR systems aligned to promote and reward implementation, and (b) an effective communication plan. Generally speaking, an effective communication plan will outline and clarify how to communicate objectives, scope, progress, and plans for implementation and training in ways that all stakeholders can understand and recognize the importance of forthcoming changes. As with any technological solution, there are multiple approaches to deployment (e.g., phased roll out, full scale) and no approach is intrinsically more effective than others. Ultimately, what approach will work best for the Air Force will depend on a host of situational factors, including stakeholder readiness and how pressing is the need(s) for new system components and changes.

---

[30] The Air Force is currently pursuing the development and prototyping of a competency-based management system to encompass technically oriented competencies (cf. The Wexford Group, 2005).

Step 4: Sustain. The last step of the implementation road map focuses on enacting the critical support activities needed to sustain the newly implemented changes over time and facilitate future changes as needed. Chief among these activities are systematic efforts to promote and maintain life cycle and user support (e.g., briefings, training, Air Force regulations and procedural directives, and so on). Consistent with our recommendations, included under life cycle support is the formalization of processes and mechanisms (i.e., agreements, such as the CEMU) to preserve linkages between AFOATS-affiliated selection systems and the Air Force's competency (performance) requirements. Finally, and most importantly, this step involves planning for and conducting regular evaluations of officer selection as a means to monitor and sustain system effectiveness.

Recommended Research to Support Road Map

To support implementation, research to answer the following questions is recommended.

How Best to Define and Assess Officer "Success"? Tracking the overall effectiveness of the Air Force's officer selection systems and estimating the ROI from different alternative HR strategies ultimately depends on how one defines and assesses officer ―success." For example, strategies that maximize (overall) training performance may not maximize actual on-the-job performance and vice versa. Similarly, those strategies that function best for maximizing individual outcomes (e.g., individuals with strong technical expertise and performance) could potentially work against the maximization of Air Force-wide outcomes. Equally as important is the choice of assessment, since some methods are more effective at providing reliable and valid information on these outcomes than others. Understanding these implications requires a systems-level (or multilevel) view uncommon to most personnel selection research, military or civilian (e.g., Ployhart, 2004; Ployhart & Schneider, 2002; Sacco & Schmitt, 2005; Schneider, Smith, & Sipe, 2000). Research in this vein would seek to (a) model the linkages (and contributions) of different, valued outcomes across multiple organizational levels, (b) identify how best to effectively and efficiently measure these different outcomes, and (c) produce reliable metrics and benchmarks that are credible with stakeholders and facilitate decision-making. An anticipated product would be a workable, model relating different, valued outcomes across multiple organizational levels so that the tradeoffs associated with different alternative HR strategies (e.g., training vs. selection) could be meaningfully and measurably compared on metrics that make sense to Air Force decision-makers.

What Critical Leader Attributes to Assess and How Best to Do So as an Integrated System? Conceptual and empirical efforts to identify what attributes are most critical to effective organizational leadership have been extensive. We made a reasonable attempt at modeling those attributes here, specifically as they apply to junior officers. At present empirical research in this area has been uneven. As a result, we know a lot more about certain critical attributes than others. Similarly, information is mixed on which assessment methods work best for measuring specific attributes. This is especially true for those critical attributes, such as cross-functional skills (e.g., social-interpersonal, self-management, etc.), that could prove the most profitable but have historically proven difficult to implement operationally because specifying what exactly said attributes are and/or developing reliable and valid methods to assess them has been problematic. Most importantly, we know relatively little about how best to combine multiple assessments into a –toolkit” – a comprehensive assessment system – that effectively and efficiently maximizes the ROI of a range of HR functions. To provide answers to these issues, research would (a) continue to systematically compare how well different assessment methods measure targeted attributes, (b) identify what methods are the most effective and efficient, (c) at what points in a junior officer's entry into the Air Force can the administration of these assessments be most beneficial to both the prospective officer and the Air Force, (d) how can the assessments of officer potential most effectively be integrated into commissioning and training (i.e., what assessments are best for officer development), and (e) which methods are best for making operational versus developmental (i.e., diagnostic) decisions or both. An anticipated product would be a well-formulated assessment –toolkit” (i.e., integrative assessment system), such as that proposed earlier, that not only selects but also facilitates the recruitment and training of junior officers.

What Would an AFOQT Form T Look Like? The current effort recommends continuing the AFOQT but, in line with recent revisions, increasing its ROI by refocusing it as a tool for broadly assessing officer potential, rather than a multiple aptitude battery that for non-rated officers mostly measures GMA. Therefore, comparable to graduate-level entry exams (e.g., GRE, LSAT, GMAT), future versions of the AFOQT would emphasize (a) higher-level cognitive aptitudes-abilities, and (b) basic cross-functional skills and personal qualities predictive of officer success (commissioning and post-commissioning). As discussed, this carries several implications for the content of future AFOQT forms. Specifically, SAT scores could be used as a substitute for assessing GMA and that in place of the affected subtests new ones be included, such as a situational judgment test (SJT) assessing several important cross-functional skills (e.g., social skills). As the SAT recently underwent fairly substantive revisions, data were not available to sufficiently and comprehensively measure some of the recommended tradeoffs. Future research in this area would (a) estimate the predictive efficacy of the new SAT, relative to AFOQT (specifically the AA subtests), across different commissioning and post-commissioning criteria (for both AFROTC and OTS), (b) prototype and pilot subtests (e.g., SJT) that could be included in an AFOQT (Form T), (c) compare the relative predictive efficacy of the new subtests (individually and in combination) to the current AFOQT (Form S), and (d) propose composites that could be employed by AFROTC and OTS decision-makers, and that preferably are aligned with competency (performance) requirements. Anticipated products include prototypes and test specifications for new subtests for future versions of the AFOQT plus composites and other tools needed to support its implementation.

How Resistant is the Self-Description Inventory (SDI+) to Compromise? While personality has been shown to be predictive of job and training performance across a wide range of jobs, including who is most likely to become a leader and perform effectively in a leadership role (Hogan & Kaiser, 2005; Judge et al., 2002; Kirkpatrick & Locke, 1991; Yukl, 2005), using self-report measures of personality in a military selection and classification context has historically proven challenging, as evident from the Army's experiences with the ABLE (e.g., White et al., 2001). Chief among these challenges is the susceptibility of self-reports to test compromise – specifically, faking and coaching effects – and the concerns these issues raise with test-takers and organizational decision-makers. Initial research with the SDI+ indicates it has promise as a screening and selection tool (e.g., Christal et al., 1997) and potentially for officer classification. As with ABLE and other self-reports, however, its long-term ROI could depend substantially on its resistance to test compromise in an actual high-stakes selection context, such as officer selection, and how it is used in that context (e.g., to rank order candidates versus an initial hurdle to screen out unqualified candidates). Research addressing this issue would (a) assess the substantive and practical impacts of possible faking and coaching effects on the SDI+ and its criterion-related validity, (b) identify best practices to prevent and minimize faking and coaching effects (e.g., issuing a warning about response verification, including a ―lie” scale, using profiles as opposed to individual dimension scores in decision-making), (c) examine Air Force stakeholders' (e.g., prospective candidates, AFROTC and AFRS decision-makers) attitudes towards the SDI+ (i.e., do stakeholders perceive the inventory as ―fakeable” and/or ―coachable”) and how these attitudes could influence its use as a selection and/or classification tool, and (d) identify and compare the potential ROI from using alternative methods for assessing personality that are presumably less susceptible to test compromise. Anticipated products would include strategies and specific tools for effectively dealing with test compromise issues and its practical implications for the use of the SDI+ in officer selection and classification.

References

Aamodt, M. G., & Williams, F. (2005). *Reliability, validity, and adverse impact of references and letters of recommendation*. Paper presented at the 20th annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Alliger, G. M., & Dwight, S. A. (2000). A meta-analytic investigation of the susceptibility of integrity tests to coaching and faking. *Educational and Psychological Measurement*, *60*, 59-72.

ANSER (2005). *Developing leaders of character at the U.S. Air Force Academy: From "first contact" to commissioning*. Colorado Springs, CO: Author.

Arth, T. O. (1986). *Validation of the AFOQT for non-rated officers* (AFHRL-TP-85-50). Brooks AFB, TX: Air Force Human Resources Laboratory.

Avolio, B. J., & Bass, B. M. (Eds.) (2002). *Developing potential across a full range of leadership: Cases in transactional and transformational leadership*. Mahwah, NJ: Lawrence Erlbaum Associates.

Avolio, B. J., Sosik, J. J., Jung, D. I., & Berson, Y. (2003). Leadership models, methods, and applications. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology, Vol. 12* (pp. 277-307). New York: John Wiley & Sons, Inc.

Bachman, J. G., Segal, D. R., Freedman-Doan, P., & O'Malley, P. M. (2000). Who chooses military service? Correlates of propensity and enlistment in the U.S. Armed Forces. *Military Psychology*, *12*, 1-30.

Baisden, A. G., Schultz, K., & Lewis-Brown, S. (2004). *A new era in U.S. Navy testing: Multimedia Navy enlisted exam*. Paper presented at the 45th Annual International Military Testing Association Conference, Brussels, Belgium.

Barrick, M. R., Day, D. V., Lord, R. G., & Alexander, R. A. (1991). Assessing the utility of executive leadership. *Leadership Quarterly*, *2*, 9-22.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, *81*, 261-272.

Bass, B. M., Avolio, B. J., Jung, D. I., & Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *Journal of Applied Psychology*, *88*, 207-218.

Breland, H. M., Kubota, M. Y., Nickerson, K., Trapani, C., & Walker, M. (2004). *New SAT writing prompt study: Analyses of group impact and reliability* (RR-2004-1). New York: The College Board.

Borman, W. C., & Brush, D. H. (1993). Toward a taxonomy of managerial performance requirements. *Human Performance*, *6*, 1-21.

Borman, W. C., Hanson, M. A., Oppler, S. H., Pulakos, E. D., & White, L. A. (1993). Role of early supervisory experience in supervisor performance. *Journal of Applied Psychology*, *78*, 443-449.

Borman, W. C., Hedge, J. W., Ferstl, K. L., Kaufman, J. D., Farmer, W. L., & Bearden, R. M. (2003). Current directions and issues in personnel selection and classification. *Research in Personnel and Human Resource Management*, *22*, 287-355.

Borman, W. C., White, L. A., Pulakos, E. D., & Oppler, S. H. (1991). Models of supervisory job performance ratings. *Journal of Applied Psychology*, *76*, 863-872.

Bowden, T. B., Keenan, P., Knapp, D. J., & Heffner, T. S. (2004). *Creating the U.S. Army Interpersonal Skills Assessment (AISA) Battery*. Paper presented at the 45th Annual International Military Testing Association (IMTA) Conference, Brussels, Belgium.

Brazell, J., & Sharon, D. (2004). *Millennial mind: Challenges of changing demographics in the applicant pool*.

Campbell, J. P. (2003, February). *Discussant's comments*. Remarks made at the U.S. Army Research Institute Workshops on New Directions in Selection & Classification, Alexandria, VA

Campbell, J. P., & Knapp, D. J. (Eds.) (2001). *Exploring the limits in personnel selection and classification*. Mahwah, NJ: Lawrence Erlbaum Associates.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.

Carretta, T. R. (2000). U.S. Air Force pilot selection and training methods. *Aviation, Space, and Environmental Medicine*, *71*, 950-956.

Carretta, T. R., & Ree, M. J. (1995). Air Force Officer Qualifying Test validity for predicting pilot training performance. *Journal of Business and Psychology*, *9*, 379-388.

Carretta, T. R., & Ree, M. J. (1996). Factor structure of the Air Force Officer Qualifying Test: Analysis and comparison. *Military Psychology*, *8*, 29-42.

Carretta, T. R., & Ree, M. J. (1998). *Near identity of cognitive structure in sex and ethnic groups* (AFHRL-TP-97-06). Brooks AFB, TX: Air Force Human Resources Laboratory.

Carretta, T. R., Retzlaff, P. D., Callister, J. D., & King, R. E. (1997). *A tale of two test batteries: A comparison of the Air Force Officer Qualifying Test and the Multdimensional Aptitude Battery* (AL/HR-TP-1997-0052). Williams AFB, AZ: Human Resources Directorate, Armstrong Laboratory.

Chan, K. Y., & Drasgow, F. (2001). Toward a theory of individual differences and leadership: Understanding the motivation to lead. *Journal of Applied Psychology*, *86*, 481-498.

Christal, R., Barucky, J. M., Driskell, W. E., & Collis, J. M. (1997). *The Air Force Self-Description Inventory (AFSDI): A summary of continuing research*. San Antonio, TX: Metrica, Inc.

Connelly, M. S., Gilbert, J. A., Zaccaro, S. J., Threlfall, K. V., Marks, M. A., & Mumford, M. D. (2000). Exploring the relationship of leadership skills and knowledge to leader performance. *Leadership Quarterly*, *11*, 65-86.

Conway, J. M. (2000). Managerial performance development constructs and personality correlates. *Human Performance*, *13*, 23-46.

Cowan, D. K., Barrett, L. E., & Wegner, T. G. (1989). *Air Force Reserve Officer Training Corps selection system validation* (AFHRL-TR-89-65). Brooks AFB, TX: Air Force Human Resources Laboratory.

Cowan, D. K., Barrett, L. E., & Wegner, T. G. (1990). *Air Force Officer Training School selection system validation* (AFHRL-TR-89-65). Brooks AFB, TX: Air Force Human Resources Laboratory.

Davison, M. L., & Davenport, E. C. (2002). Identifying criterion-related patterns of predictor scores using multiple regression. *Psychological Methods*, *7*, 468-484.

Day, D. V. & Lord, R. G. (1988). Executive leadership and organizational performance: Suggestions for a new theory and methodology. *Journal of Management*, *14*, 453-464.

Day, D. V., Shleicher, D. J., Unckless, A. L., & Hiller, N. J. (2002). Self-monitoring personality at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology*, *87*, 390-401.

Drasgow, F. (2003). Intelligence and the workplace. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology, Vol. 12* (pp. 107-130). New York: John Wiley & Sons, Inc.

Diehl, G. E. (1986). *Correlations among SAT, ACT, AFOQT, and grade point average* (AD-A190 251). Maxwell AFB, AL: Air University.

Earles, J. A., & Ree, M. J. (1991). *Air Force Officer Qualifying Test (AFOQT): Estimating the general ability component* (AFHRL-TP-91-39). Brooks AFB, TX: Air Force Human Resources Laboratory.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology*, *84*, 155-166.

Fallesen, J. J., & Reichard, R. J. (2005). *Leadership competencies: Building a foundation for Army leader development.* Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Ford, L. A., Campbell, R. C., Campbell, J. P., Knapp, D. J., & Walker, C. B. (2000). *21st century soldiers and noncommissioned officers: Critical predictors of performance* (TR-1102). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Gallagher, C. A., Joseph, L. E., & Park, M. V. (2002). Implementing organizational change. In J. W. Hedge & E. D. Pulakos (Eds.), *Implementing organizational interventions: Steps, processes, and best practices* (pp. 12-42). San Francisco: Jossey-Bass.

Galway, L. A., Buddin, R., Thirtle, M. R., Ellis, P., & Mele, J. D. (2005). *Understrength Air Force officer career fields: A force management approach* (MG-131). Santa Monica, CA: RAND Corporation.

Garstka, J. J. (2003, October 1). *Integrating innovation, leadership, and cultural change*. Presentation made at the Workshop on Transforming the Culture of DoD.

Geraghty, B. J., & Collins, J. M. (2003). Becoming brass: Task-related and contextual components of military officer jobs. *Military Psychology*, *15*, 255-272.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology*, *87*, 268-279.

Hartke, D. D., & Short, L. O. (1988). *Validity of the Academic Aptitude composite of the Air Force Officer Qualifying Test (AFOQT)* (AFHRL-TP-87-61). Brooks AFB, TX: Air Force Human Resources Laboratory.

Hedge, J. W., Borman, W. C., Bruskiewicz, K. T., & Bourne, M. J. (2004). The development of an integrated performance category system for supervisory jobs in the U.S. Navy. *Military Psychology*, *16*, 231-243.

Hedge, J. W., & Teachout, M. S. (1986). *Job performance measurement: A systematic program of research and development* (AFHRL-TP-86-37). Brooks AFB: Air Force Human Resources Lab.

Higgs, A. C., Papper, E. M., & Carr, L. S. (2000). Integrating selection with other organizational processes and systems. In J. F. Kehoe (Ed.), *Managing selection in changing organizations: Human resource strategies* (pp. 73-122). San Francisco: Jossey-Bass.

Hogan, R., & Kaiser, R. B. (2005). What we know about leadership. *Review of General Psychology*, *9*, 169-180.

Hooper, A. C., Cullen, M. J., & Sackett, P. R. (2006). Operational threats to the use of SJTs: Faking, coaching, and retesting issues. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 205-232). Lawrence Erlbaum Associates.

Horey, J. D., & Fallesen, J. J. (2003). *Leadership competencies: Are we all saying the same thing?* Paper presented at the 45th Annual Conference of the International Military Testing Association, Pensacola, FL.

Horey, J. D., & Fallesen, J. J. (2004). *Leadership competencies for contemporary Army operations: Development, review and validation*. Paper presented at the 46th Annual Conference of the International Military Testing Association, Brussels, Belgium.

Horey, J. D., Fallesen, J. J., Morath, R., Cronin, B., Cassella, R., Franks, Jr., W., & Smith, J. (2004). *Competency-based future leadership requirements* (TR-05-1148). Alexandria, VA: U.S. Army Research Institute.

Horey, J., Morath, R., McGonigle, T., Cronin, B., & Wilson, S. (2005). *Air Force strategic initiatives research task D: Force quality report*. Booz Allen Hamilton, McClean, VA.

Hossek, J. R., Mattock, M. G., Fair, C. C., Sharp, J., & Totten, M. (2004). *Attracting the best: How the military competes for information technology personnel* (MG-108-OSD). Santa Monica, CA: RAND Corporation.

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance*, *11*, 209-244.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, *75*, 581-595.

Hough, L. M., & Ones, D. S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology* (Vol. 1, pp. 233-277). London: Sage.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment*, *9*, 152-194.

Hunter, J. E. (1983). *The prediction of job performance in the military using ability composites: The dominance of general cognitive ability over specific aptitudes*. Rockville, MD: Research Applications.

Ilgen, D. R., & Pulakos, E. D. (Eds.) (1999). *The changing nature of performance: Implications for staffing, motivation, and development*. San Francisco: Jossey-Bass.

Jones, J. W., & Dages, K. D. (2003). Technology trends in staffing and assessment: A practice note. *International Journal of Selection and Assessment*, *11*, 247-252.

Jones, R. G., Stevens, M. J., & Fischer, D. L. (2000). Selection in team contexts. In J. F. Kehoe (Ed.), *Managing selection in changing organizations: Human resource strategies* (pp. 210-241). San Francisco: Jossey-Bass.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*, 765-780.

Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, *89*, 542-552.

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, *89*, 755-768.

Judge, T. A., Piccolo, R. F., & Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology*, *89*, 36-51.

Kilburn, M. R., & Asch, B. J. (2003). *Recruiting youth in the college market: Current practices and future policy options* (MR-1093-OSD). Santa Monica, CA: RAND Corporation.

Kirkpatrick, S. A., & Locke, E. A. (1991). Leadership: Do traits matter? *Academy of Management Executive*, *5*, 48-60.

Kobrin, J. L., & Schmidt, A. E. (2005). *The research behind the new SAT* (RS-11). New York: The College Board.

Knapp, D. J., & Campbell, R. C. (Eds.) (2005). *Army enlisted competency assessment program: Phase II report* (FR-05-06). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Knapp, D. J., McCloy, R. A., & Heffner, T. S. (Eds.) (2004). *Validation of measures designed to maximize 21st century Army NCO performance* (TR-1145). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Lewis, P. M. (1995). *Conceptual complexity and officer effectiveness*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

McDaniel, M. A., Morgeson, F. P., Bruhn Finnegan, E., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, *86*, 730-740.

McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, *85*, 812-821.

McIntyre, R. M., Jordan, P., Mergen, C., Hamill, L., & Jacobs, T. O. (1993). *The construct validity of the CPA: Report on three investigations* (TR-92-017). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Michael, P. G., Chen, H. T., Janega, J. B., Farmer, W. L., Eller, E. D., & Nayak, A. (2005). *Validity evidence for a computer-administered interest inventory for the Navy*. Paper presented at the 20th annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Michaels, E., Handfield-Jones, H., & Axelrod, B. (2001). *The war for talent*. Boston: Harvard Business School Press.

Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, *55*, 1-22.

Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology*, *49*, 229-258.

Motowidlo, S. J. (2003). Job performance. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology, Vol. 12* (pp. 39-53). New York: John Wiley & Sons, Inc.

Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology*, *79*, 475-480.

Mueller-Hanson, R., Heggestad, E. D., & Thornton, III, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, *88*, 348-355.

Murphy, K. R., & Shiarella, A. H. (1997). Implications of the multidimensional nature of job performance for the validity of selection tests: Multivariate frameworks for studying test validity. *Personnel Psychology*, *50*, 823-854.

Nguyen, N. T., McDaniel, M. A., & Whetzel, D. (2005). *Subgroup differences in situational judgment test performance: A meta-analysis*. Paper presented at the 20th annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Olea, M. M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than g. *Journal of Applied Psychology*, *79*, 845-851.

Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, *11*, 245-269.

Orvis, B. R., & Asch, B. J. (2001). *Military recruiting: Trends, outlook, and implications* (MR-902-A/OSD). Santa Monica, CA: RAND Corporation.

Pearlman, K. (2003, February). *Private sector perspectives on selection content and methodology*. Presentation made at the U.S. Army Research Institute Workshops on New Directions in Selection & Classification, Alexandria, VA.

Pearlman, K., & Barney, M. F. (2000). Selection for a changing workplace. In J. F. Kehoe (Ed.), *Managing selection in changing organizations: Human resource strategies* (pp. 3-72). San Francisco: Jossey-Bass.

Ployhart, R. E. (2004). Organizational staffing: A multilevel review, synthesis, and model. *Research in Personnel and Human Resources Management*, *23*, 121-176.

Ployhart, R. E., & Schneider, B. (2002). A multilevel perspective on personnel selection research and practice: Implications for selection system design, assessment, and construct validation. In F. J. Dansereau & F. Yamarino (Eds.), *Research in multi-level issues: The many faces of multi-level issues* (Vol. 1, pp. 95-140). Oxford, UK: Elsevier.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology*, *85*, 612-624.

Pulakos, E. D., Schmitt, N., Dorsey, D. W., Arad, S., Hedge, J. W., & Borman, W. C. (2002). Predicting adaptive performance: Further tests of a model of adaptability. *Human Performance*, *15*, 299-324.

Ree, M. J., & Carretta, T. R. (1998). *Interchangeability of verbal and quantitative scores for personnel selection: An example* (AL/HR-TP-1997-0016). Mesa, AZ: Air Force Research Laboratory.

Ree, M. J., Carretta, T. R., & Earles, J. A. (2003). Salvaging construct equivalence through equating. *Personality and Individual Differences*, *35*, 1293-1305.

Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology*, *80*, 721-780.

Roberts, H. E., & Skinner, J. (1996). Gender and racial equity of the Air Force Officer Qualifying Test in officer training school selection decisions. *Military Psychology*, *8*, 95-113.

Rosse, J. G., Stecher, M. D, Miller, J. L. & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, *83*, 634-644.

Sacco, J. M., & Schmitt, N. (2005). Dynamic multilevel model of demographic diversity and misfit effects. *Journal of Applied Psychology*, *90*, 203-231.

Sackett, P. R., Berry, C. M., & Wiemann, S. A. (2005). *A review of recent developments in integrity test research*. Paper presented at the 20th annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Sackett, P. R., & Mavor, A. S. (2003). *Attitudes, aptitudes, and aspirations of American youth: Implications for military recruitment*. Washington, DC: National Academies Press.

Sager, C. E., Russell, T. L., Campbell, R. C., & Ford, L. A. (2004). *Future soldiers: Analysis for entry-level performance requirements and their predictors* (FR-03-82). Alexandria, VA: Human Resources Research Organization.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.

Shore, C. W, Gould, R. B., Ree, M. J., Alley, W. E., & Skinner, J. (2003). *Air Force Officer Qualifying Test (AFOQT): Form R and S development and norms*. Unpublished technical report.

Sperl, T. C., Ree, M. J., & Steuck, K. W. (1992). Armed Services Vocational Aptitude Battery and Air Force Officer Qualifying Test: Analyses of common attributes. *Military Psychology*, *4*, 175-188.

Stark, S., Chernyshenko, O. S., Chan, K. Y., Lee, W. C., & Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology*, *86*, 943-953.

U.S. Department of the Air Force. *USAFA candidate evaluation/AFROTC college scholarship evaluation* (AF Form 4060). Washington, DC: Author.

U.S. Department of the Air Force (1996). *Global engagement: A vision for the 21st century Air Force*. Washington, DC: Author.

U.S. Department of the Air Force (1999, October 1). *Evaluating USAF Academy (USAFA) candidates and Air Force Reserve Officer Training Corps (AFROTC) college scholarship applicants* (AFI 36-2249). Washington, DC: Author.

U.S. Department of the Air Force (2001, February 3). *AFROTC POC, pilot and navigator allocations management* (AFOATSI 36-2013). Washington, DC: Author.

U.S. Department of the Air Force (2003). *Personnel strategic plan: FY 2004-2009*. Washington, DC: Author.

U.S. Department of the Air Force (2004a). *Leadership and force development* (AFDD 1-1). Washington, DC: Author.

U.S. Department of the Air Force (2004b). *The U.S. Air Force transformation flight plan*. Washington, DC: Author.

U.S. Department of the Air Force (2005). *U.S. Air Force posture statement 2005*. Washington, DC: Author.

U.S. Department of the Army (1999). *Army leadership: Be, know, do* (FM 22-100). Washington, DC: Author.

U.S. Department of Defense (2001, September). *2001 Quadrennial Defense Review report*. Washington, DC: Author.

U.S. Department of Defense (2005). *June 2005 Youth Poll 9 briefing*. Retrieved November 14, 2005 from http://www.dmren.org/DMREN/execute/secure/documents/abstract?id=313

U.S. General Accountability Office (2005). *Military personnel: Reporting additional servicemember demographics could enhance congressional oversight* (GAO-05-952). Washington, DC: Author.

U.S. Joint Chiefs of Staff (2000). *Joint vision 2020*. Washington, DC: U.S. Department of Defense.

U.S. Joint Chiefs of Staff (2003). *An evolving joint perspective: U.S. joint warfare and crisis resolution in the 21st century*. Washington, DC: U.S. Department of Defense.

U.S. Office of the Secretary of Defense (2002). *Population representation in the military services FY 2001*. Washington, DC: Author.

U.S. Office of the Secretary of Defense (2003). *Population representation in the military services FY 2002*. Washington, DC: Author.

U.S. Office of the Secretary of Defense (2004). *Population representation in the military services FY 2003*. Washington, DC: Author.

U.S. Office of the Secretary of Defense (2005). *Population representation in the military services FY 2004*. Washington, DC: Author.

Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology*, *81*, 525-531.

Vasilopoulos, N. L., Reilly, R. R., & Leaman, J. A. (2000). The influence of job familiarity and impression management on self-report measure scale scores and response latencies. *Journal of Applied Psychology*, *85*, 50-64.

Viswesvaran, C. & Ones, D. S. (1999). Meta-analysis of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, *59*, 197-210.

Wai, J., Lubinski, D., & Benbow, C. P. (2005). Creativity and occupational accomplishments among intellectually precocious youths: An age 13 to age 33 longitudinal study. *Journal of Educational Psychology*, *97*, 484-492.

Weaver, N. (2001). *Developing aerospace leaders for the 21st century* (P-8060). Santa Monica, CA: RAND Corporation.

Weeks, J. L. (2000). *USAF pilot selection* (AFRL-HE-TP-2000-04). Mesa, AZ: Air Force Research Laboratory.

Weissmuller, J. J., Schwartz, K. L., Kenney, S. D., Shore, C. W, & Gould, R. B. (2004). *Recent developments in USAF officer testing and selection*. Paper presented at the 46th Annual Conference of the International Military Testing Association, Brussels, Belgium.

Welsh, J. R., Kucinkas, S. K., & Curran, L. T. (1990). *Armed Services Vocational Battery (ASVAB): Integrative review of validity studies* (AFHRL-90-22). Brooks AFB, TX: Air Force Human Resources Laboratory.

Wexford Group International (2005). *The coordinated Air Force Competency-Based Management System (CBMS) implementation plan*. Arlington, VA: Air Force Directorate of Personnel.

White, L. A., Young, M. C., & Rumsey, M. G. (2001). ABLE implementation issues and related research. In J. P. Campbell & D. J. Knapp (Eds.), *Exploring the limits in personnel selection and classification* (pp. 525-558). Mahwah, NJ: Lawrence Erlbaum Associates.

Wong, L., Bliese, P., & McGurk, D. (2003). Military leadership: A context specific review. *Leadership Quarterly*, *14*, 657-692.

Wright, P. M., & Boswell, W. R. (2002). Desegregating HRM: A review and synthesis of micro and macro HR research. *Journal of Management*, *28*, 247-276.

Yukl, G. (2005). *Leadership in organizations* (6th ed.). Englewood Cliffs, NJ: Prentice Hall.

Zaccaro, S. J. (1999). Social complexity and the competencies required for effective military leadership. In J. G. Hunt, G. E. Dodge, & L. Wong (Eds.), *Out-of-the-box leadership: Transforming the twenty-first century Army and other top performing organizations* (pp. 131-151). Stamford, CT: JAI Press.

Zaccaro, S. J. (2001). *The nature of executive leadership: A conceptual and empirical analysis of success*. Washington, DC: American Psychological Association.

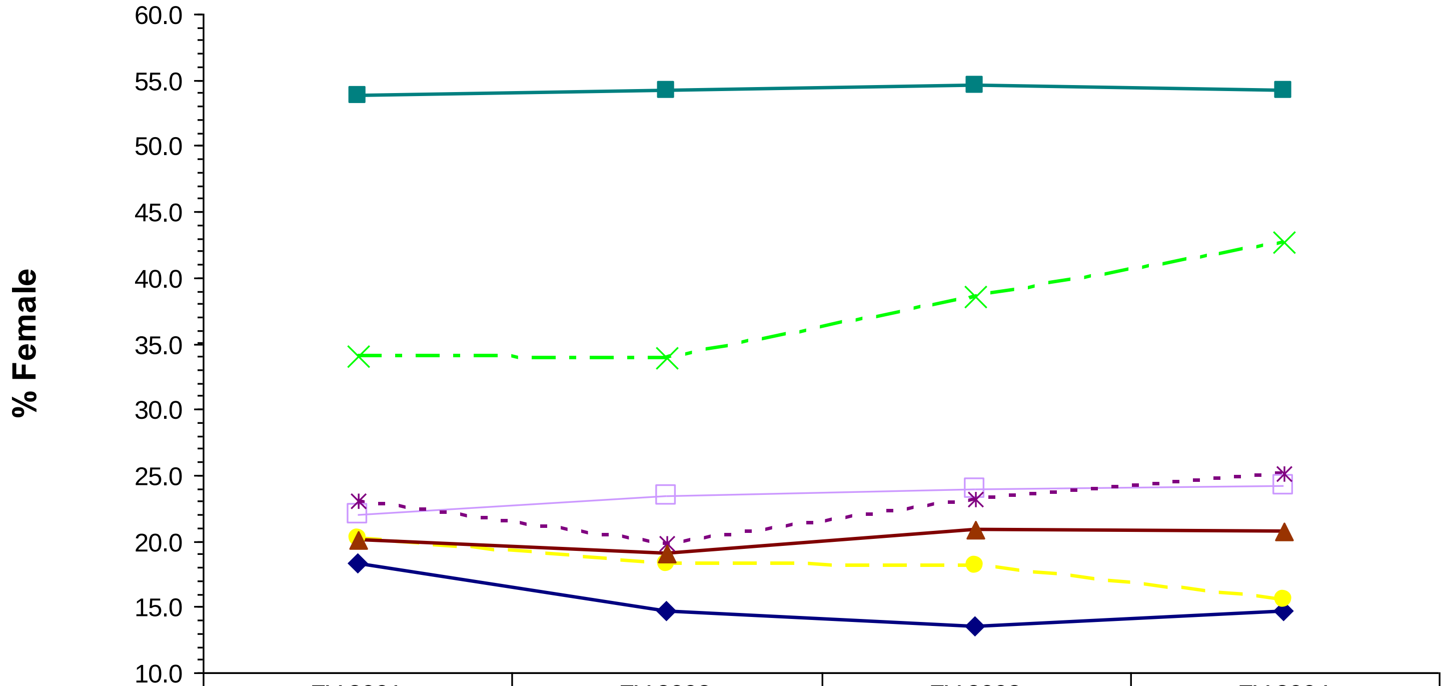
Zaccaro, S. J., Kemp, C., & Bader, P. (2004). Leader traits and attributes. In J. Antonakis, A.T. Cianciolo, & R.J. Sternberg (Eds.), *The nature of leadership* (pp. 101-124). Thousand Oaks, CA: Sage.

Zaccaro, S. J., & Klimoski, R. J. (Eds.) (2001). *The nature of organizational leadership: Understanding the imperatives confronting today's leaders*. San Francisco: Jossey-Bass.

Zaccaro, S. J., Klimoski, R. J., & Boyce, L. A. (1999). *The changing U.S. Army: A summary of future focused reports from 1990-1999* (RR-99-1747). Alexandria, VA: U.S. Army Research Institute (ARI).

Appendix A:
Demographic Composition of Air Force (Active Component) Officer Accessions by Source of Commissioning for FYs 2001-2004

Figure A-1. Percentage of Female (Active Component) Officer Accessions by Source of Commissioning with Total DoD and Civilian Comparison Group (FYs 2001-2004)



	FY 2001	FY 2002	FY 2003	FY 2004
USAFA	18.24	14.61	13.46	14.65
AFROTC	22.01	23.39	23.94	24.23
OTS	20.25	18.24	18.16	15.60
Dir Appt	33.97	33.94	38.53	42.72
Other	22.99	19.82	23.16	25.11
Total DoD	20.07	19.10	20.83	20.68
Civilian	53.84	54.21	54.58	54.14

Notes. Total DoD comparison group represents an aggregate across the Services (Army, Navy, Marine Corp, Air Force). Civilian comparison group based on 21-35 year old college graduates in the non-institutional civilian population. Civilian comparison group data from the Bureau of Labor Statistics.

Figure A-2. Number of Female (Active Component) Officer Accessions by Source of Commissioning (FYs 2001-2004)

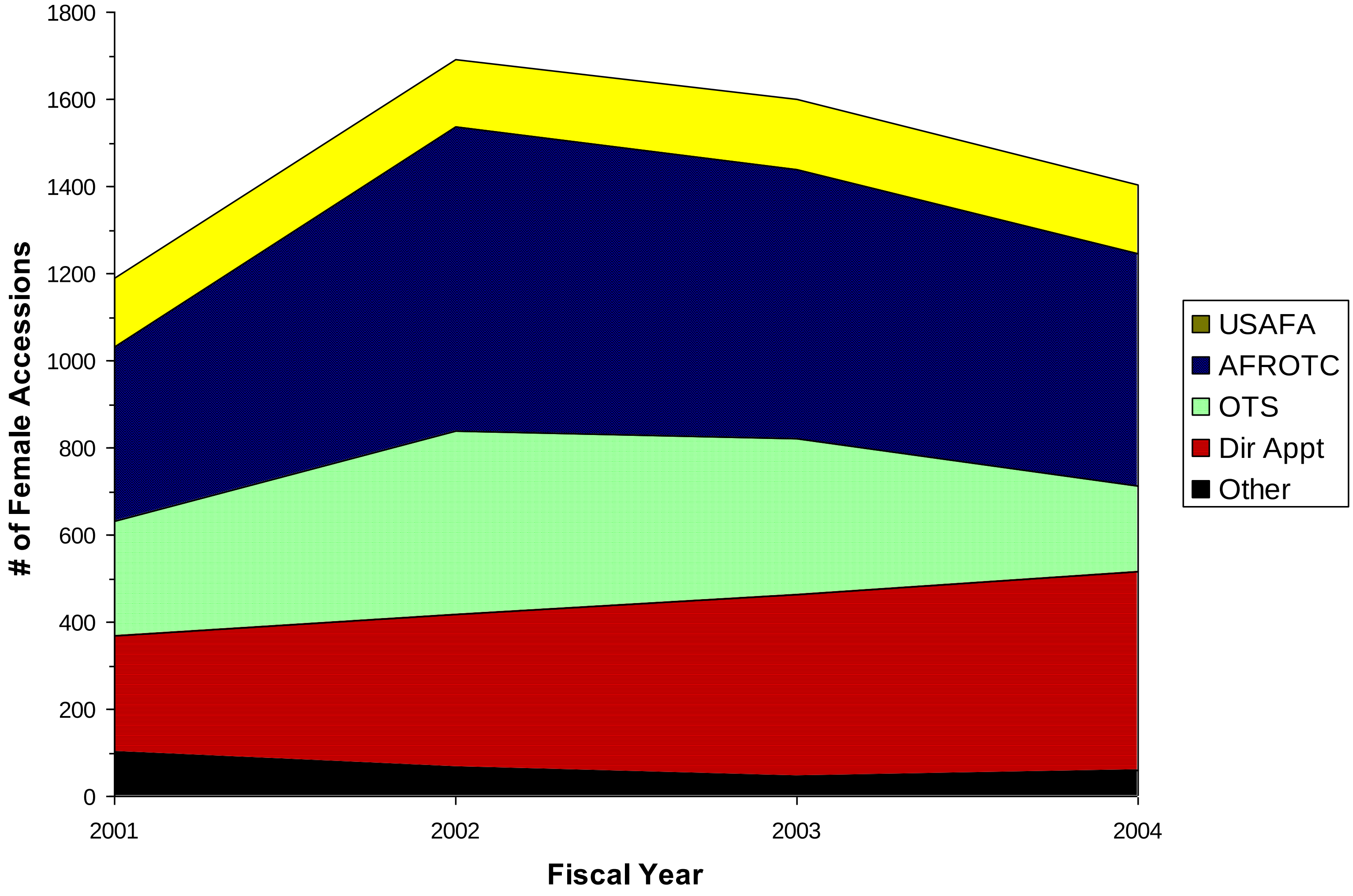
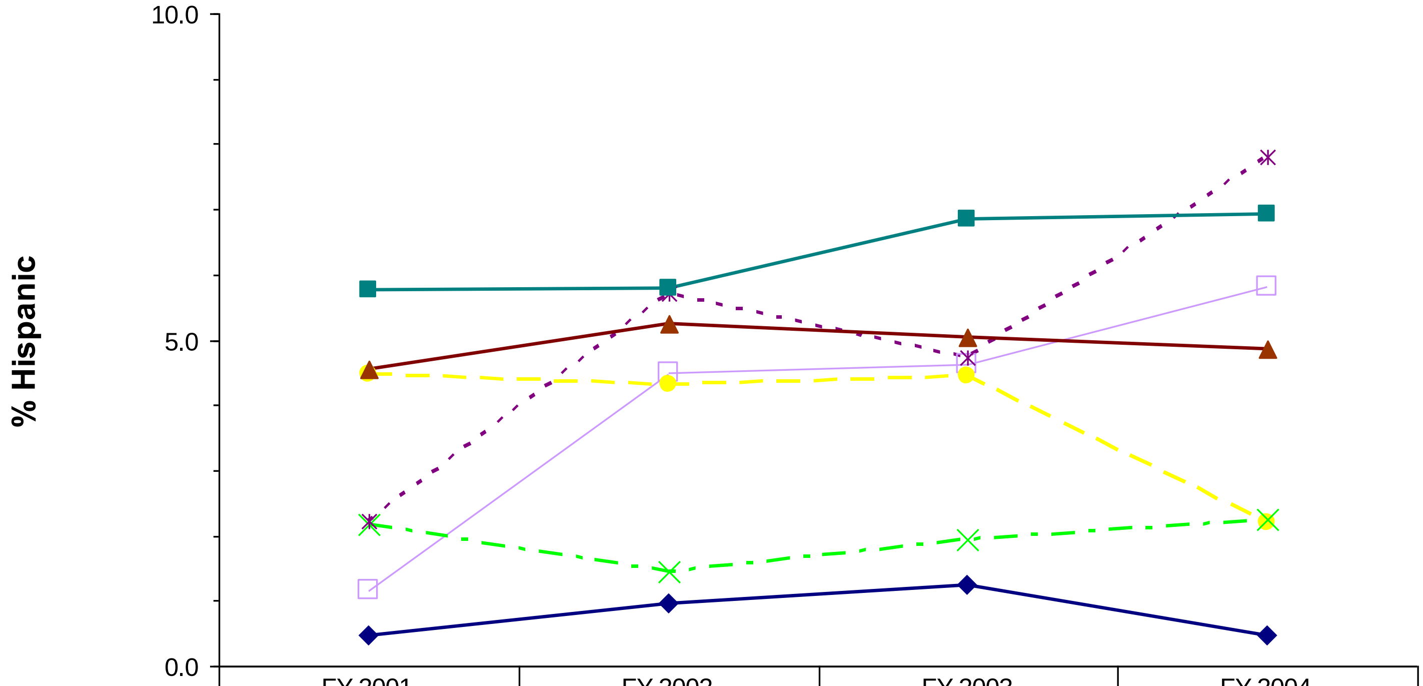


Figure A-3. Percentage of Hispanic (Active Component) Officer Accessions by Source of Commissioning with Total DoD and Civilian Comparison Group (FYs 2001-2004)



	FY 2001	FY 2002	FY 2003	FY 2004
◆ USAFA	0.46	0.95	1.24	0.47
□ AFROTC	1.16	4.50	4.62	5.82
● OTS	4.48	4.32	4.44	2.21
× Dir Appt	2.16	1.45	1.93	2.24
✳ Other	2.23	5.71	4.74	7.79
▲ Total DoD	4.54	5.24	5.03	4.87
■ Civilian	5.76	5.78	6.85	6.93

Notes. Total DoD comparison group represents an aggregate across the Services (Army, Navy, Marine Corp, Air Force). Civilian comparison group based on 21-35 year old college graduates in the non-institutional civilian population. Civilian comparison group data from the Bureau of Labor Statistics.

Figure A-4. Number of Hispanic (Active Component) Officer Accessions by Source of Commissioning (FYs 2001-2004)

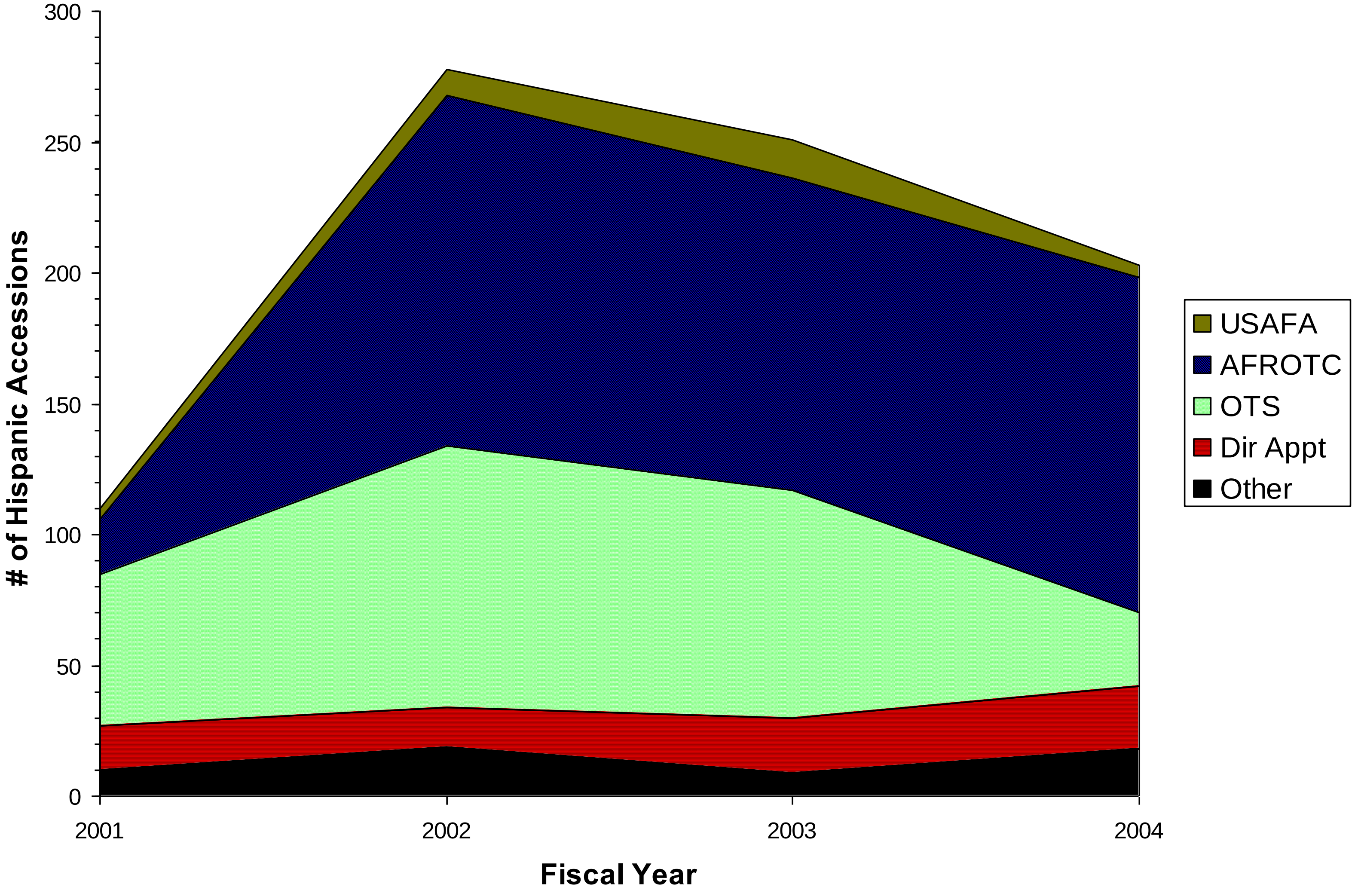
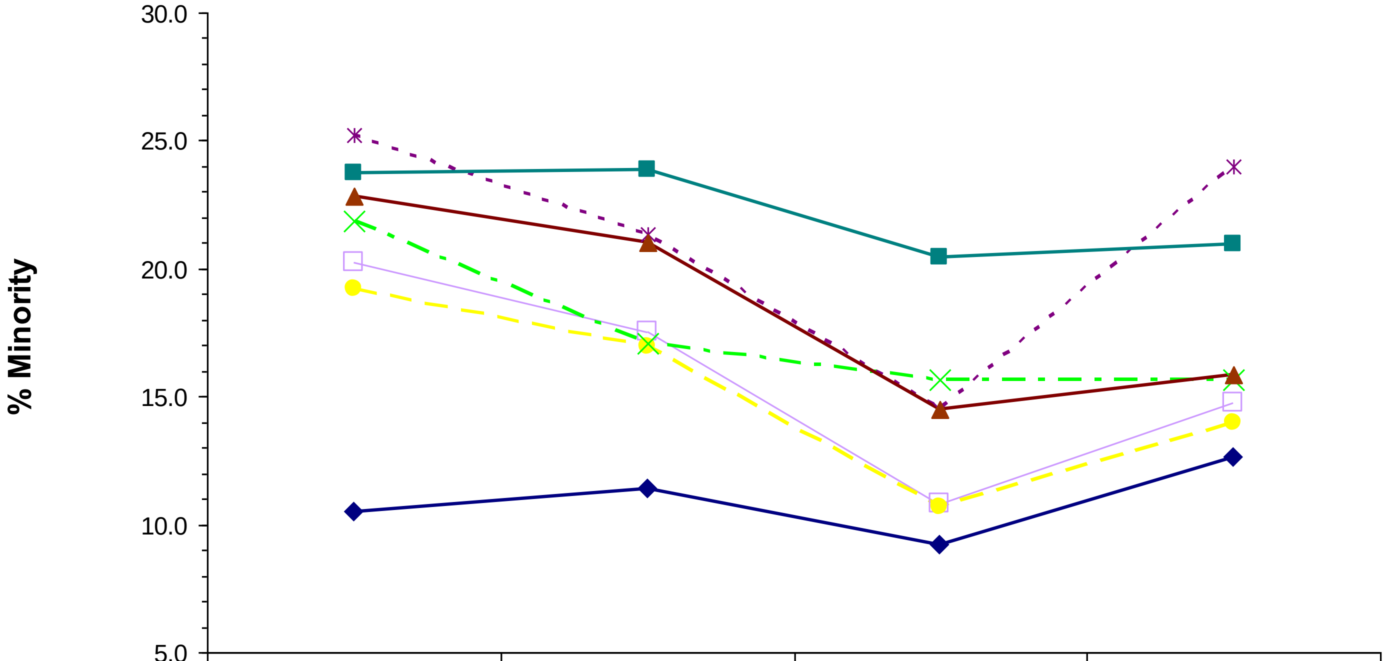


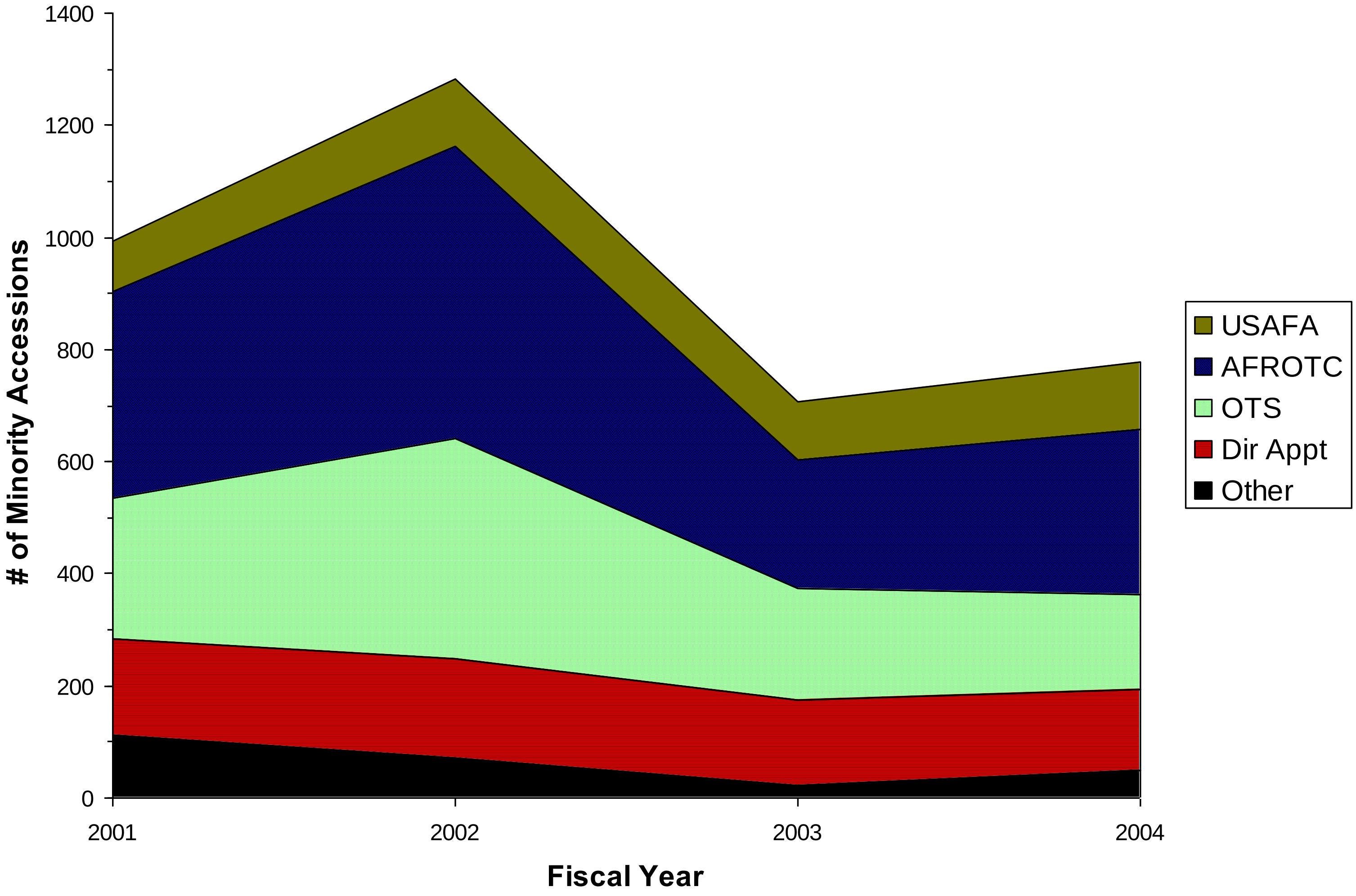
Figure A-5. Percentage of Minority (Active Component) Officer Accessions by Source of Commissioning with Total DoD and Civilian Comparison Group (FYs 2001-2004)



	FY 2001	FY 2002	FY 2003	FY 2004
USAFA	10.51	11.39	9.20	12.61
AFROTC	20.25	17.54	10.80	14.77
OTS	19.24	16.95	10.70	13.98
Dir Appt	21.88	17.07	15.65	15.68
Other	25.22	21.32	14.47	24.00
Total DoD	22.83	21.05	14.51	15.83
Civilian	23.76	23.87	20.42	20.93

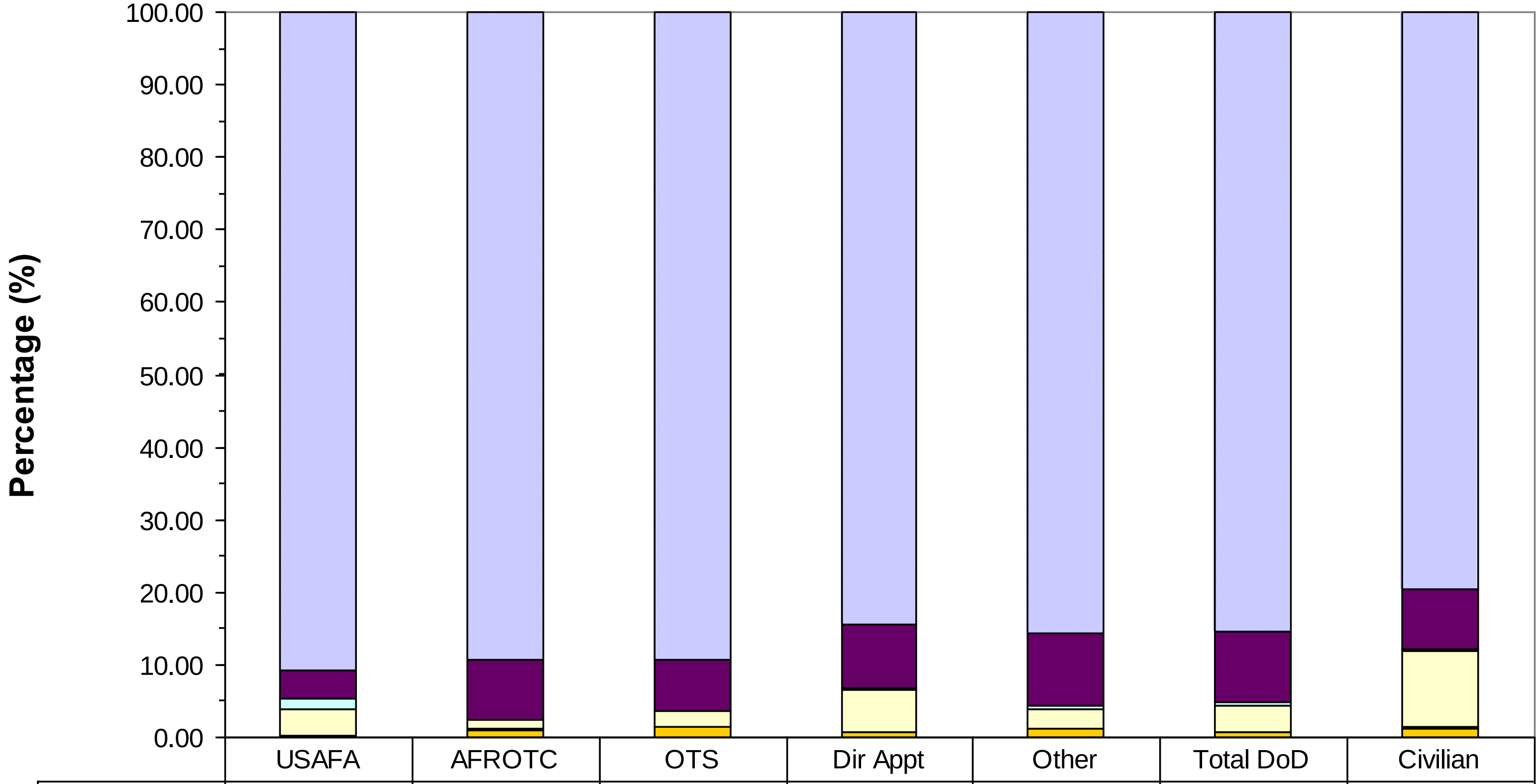
Notes. Total DoD comparison group represents an aggregate across the Services (Army, Navy, Marine Corp, Air Force). Civilian comparison group based on 21-35 year old college graduates in the non-institutional civilian population. Civilian comparison group data from the Bureau of Labor Statistics. Because of changes in how DoD codes race/ethnicity, the computation of % Minority varies across FYs. For FYs 2000-2001, % Minority represents an aggregate of Black, Hispanic, and Other. For FYs 2002-2003, % Minority represents an aggregate of Black, AIAN, Asian, NHPI, and Two or More.

Figure A-6. Number of Minority (Active Component) Officer Accessions by Source of Commissioning (FYs 2001-2004)



Notes. Because of changes in how DoD codes race/ethnicity, the computation of % Minority varies across FYs. For FYs 2000-2001, % Minority represents an aggregate of Black, Hispanic, and Other. For FYs 2002-2003, % Minority represents an aggregate of Black, AIAN, Asian, NHPI, and Two or More.

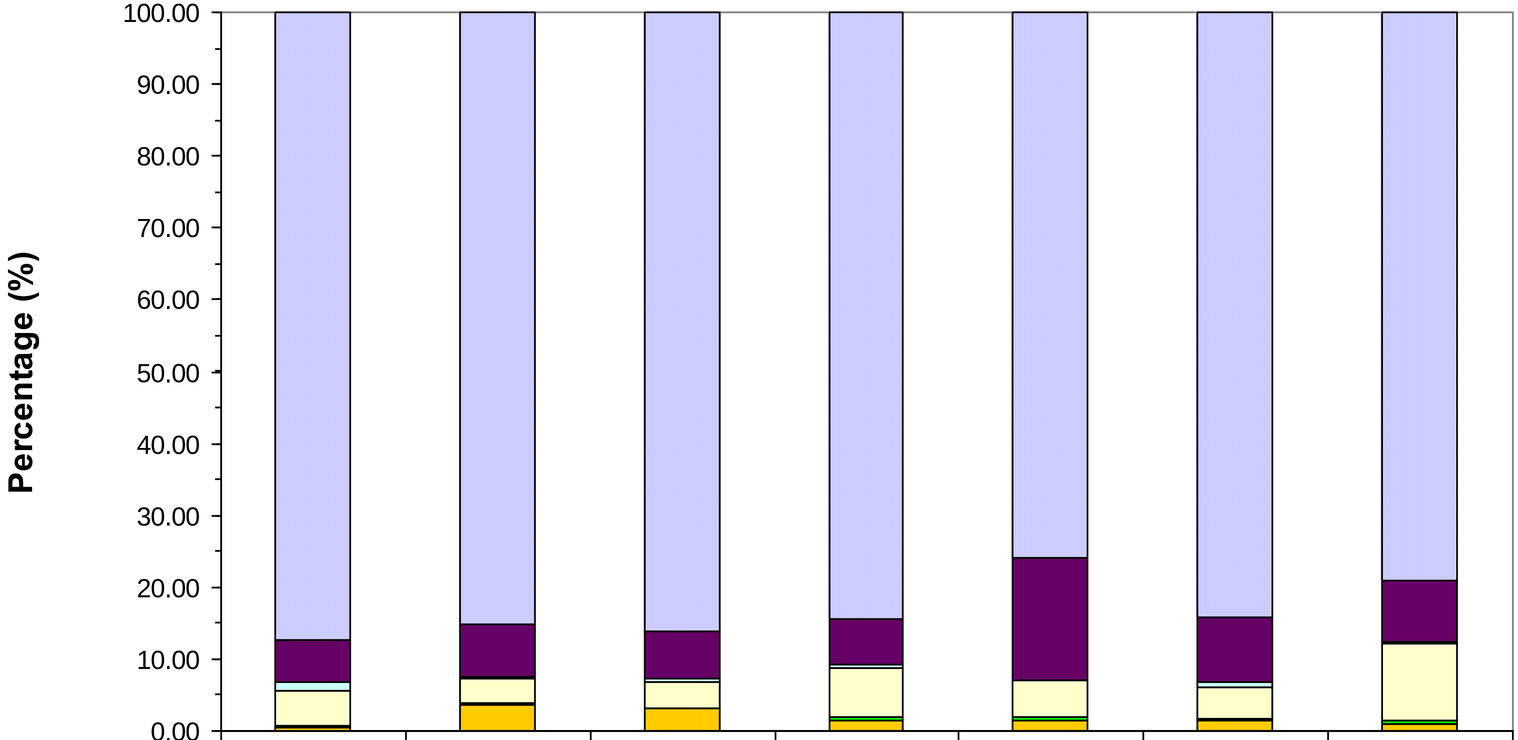
Figure A-7. Racial Composition of (Active Component) Officer Accessions by Source of Commissioning with Total DoD and Civilian Comparison Groups (FY 2003)



	USAFA	AFROTC	OTS	Dir Appt	Other	Total DoD	Civilian
White	90.80	89.20	89.30	84.35	85.53	85.49	79.58
Black	3.81	8.29	6.96	8.75	10.06	9.65	8.20
AIAN	1.59	0.09	0.21	0.41	0.63	0.39	0.33
Asian	3.45	1.28	1.98	5.77	2.52	3.72	10.43
NHPI	0.00	0.24	0.05	0.00	0.00	0.08	0.27
Two or more	0.35	0.90	1.50	0.72	1.26	0.67	1.20

Notes. AIAN = American Indian/Alaskan Native; NHPI = Native Hawaiian/Pacific Islander.

Figure A-8. Racial Composition of (Active Component) Officer Accessions by Source of Commissioning with Total DoD and Civilian Comparison Groups (FY 2004)



	USAFA	AFROTC	OTS	Dir Appt	Other	Total DoD	Civilian
☐ White	87.39	85.23	86.02	84.32	76.00	84.18	79.07
■ Black	5.88	7.23	6.79	6.49	17.00	9.11	8.44
☐ AIAN	1.16	0.25	0.33	0.43	0.00	0.53	0.36
☐ Asian	4.94	3.37	3.60	6.81	5.00	4.46	10.75
🟩 NHPI	0.11	0.25	0.16	0.54	0.50	0.17	0.31
🟨 Two or more	0.53	3.67	3.11	1.41	1.50	1.55	1.08

Notes. AIAN = American Indian/Alaskan Native; NHPI = Native Hawaiian/Pacific Islander.

Appendix B:
AFOQT and SAT-ACT Equivalency Chart

SAT-R Equivalency Chart

AFOQT AA	ACT COMP	SAT-O TOTAL	SAT-R TOTAL
		Test Date Before 1 Apr 95	Test Date On / After 1 Apr 95
10		710	840
11		720	850
12, 13		730, 740	860
14	18	750	870
		760	880
15		770	890
16		780	900
17	19	790	910
18		800	920
19		810	930
20		820	940
21, 22	20	830	950
23		840	960
24, 25		850, 860	970
26		870	980
27	21	880	990
28, 29		890	1000
30		900	1010
31		910	1020
32, 33	22	920	1030
34		930	1040
35, 36, 37		940, 950	1050
38, 39		960	1060
40	23	970	1070
41, 42		980	1080
43		990	1090
44, 45		1000	1100
46, 47	24	1010	1110
48, 49, 50		1020, 1030	1120
51, 52		1040	1130
53	25	1050	1140
54, 55		1060	1150
56, 57		1070	1160
58		1080	1170
59, 60, 61, 62	26	1090, 1100	1180
63		1110	1190
64, 65		1120	1200
66, 67		1130	1210
68	27	1140	1220
69, 70		1150	1230
71, 72		1160, 1170	1240

Source. AFOATSI 36-2013, *AFROTC POC, Pilot and Navigator Allocations Management* (p. 5)

73, 74		1180	1250
75	28	1190	1260
76, 77		1200	1270
78		1210	1280
79		1220	1290
80	29	1230	1300
81, 82, 83		1240, 1250	1310
84		1260	1320
85		1270	1330
86	30	1280	1340
87		1290	1350
88		1300	1360
89		1310	1370
90	31	1320	1380
91		1330	1390
92		1340	1400
		1350	1410
93	32	1360	1420
94		1370	1430
95		1380	1440
96		1390	1460
	33	1400	1470
97		1410	1480
98		1420	1490
		1430	1510
99	34	1440	1520
		1450	1530
	35	1460	1550
		1470	1560
		1480	1580
	36	1490-1600	1600